

MEASURING PROGRESS IN CLINICAL GOVERNANCE

By

TIMOTHY FREEMAN

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY

School of Public Policy
The University of Birmingham
June 2003

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

The role of clinical governance within the New Labour NHS reforms of 1997 is to ensure dependable local service delivery of central targets by supporting quality assurance and improvement activities. These twin objectives result in significant tensions, and health care organisations face a difficult balancing act as they seek to align central and local agendas for quality improvement while meeting performance measurement requirements.

The thesis situates clinical governance within broader governance trends, reviewing empirical evaluations of implementation and tracing the effects of internal tensions. It synthesises current literature on applying performance indicators to public sector bureaucracies, and makes a case for the development of two new validated measures of clinical governance: a measure of climate at the level of individual staff (CGCQ); and a competency framework for application at directorate and corporate levels (OPCG).

Following detailed empirical development and assessment of validity and reliability, the measures are applied: the CGCQ to a sample of 1140 participants across four case study sites; and the OPCG to a national sample of 1916 staff from acute, ambulance and mental health / social care trusts. Results indicate the dominance of the quality assurance agenda and the need to support internal quality improvement initiatives.

Dedication

This thesis would not have been completed but for the support provided by my family. To this end, I dedicate the thesis to:

My partner Sarah, for her indulgence and understanding in this matter so very far beyond what could have been reasonably expected;

My son Joshua, for placing everything in such perspective, a ready smile and for those rare nights of sleeping-through (!);

And to my parents Pat and Norman Hares, for awakening an interest in education that ultimately led to this point.

Acknowledgements

I am deeply indebted to Professor Peter Spurgeon, whose patience, guidance and unfailing good humour proved invaluable throughout my studies at the University of Birmingham. I am also grateful to the former West Midlands Regional Health Authority for their generous funding of a clinical governance fellowship, which enabled me the luxury of time to develop the measures reported in this thesis. Additional thanks are also due to the National Audit Office for funding the national Organisational Progress in Clinical Governance (OPCG) study, and to all those who gave so freely of their valuable time to complete questionnaires or take part in focus or Q-sort groups.

Table of Contents

		Page
	Volume I	
Chapter 1	Clinical governance: a critical introduction	1
	Introduction	
1.1	Clinical governance, its origins and status	4
1.1.1	Definitions	4
1.1.2	Governance and modes of governance	8
1.1.3	Principal – agent theory	13
1.1.4	Historical development of modes of governance in the UK NHS	15
1.1.5	The continued rise of the regulatory state?	26
1.2	The importance of measuring progress in clinical governance	30
1.3	The contribution of the thesis	31
	Conclusions	34
Chapter 2	The early development of clinical governance: a review of empirical evaluations	37
	Introduction	
2.1	Methods	38
2.2	Results	40
2.2.1	Primary care	40
2.2.2	Secondary and tertiary care	47
2.2.3	Shared potential future difficulties	56
	Conclusions	59
Chapter 3	Using performance indicators to improve health care quality in the public sector: a review of the literature	63
	Introduction	
3.1	The rise of performance indicators	65

		Page
3.2	Methods	66
3.3	Results	68
3.3.1	Aims of performance indicator systems	69
3.3.2	Perceived or reported problems with performance indicator systems	72
3.3.3	Perceived or reported factors facilitating the derivation, implementation and use of performance indicators	81
	Conclusions	85
Chapter 4	Measuring progress in clinical governance: current options and the case for additional measures	88
	Introduction	
4.1	Measurement options: a conceptual framework	89
4.2	External approaches	93
4.2.1	External / summative	96
4.2.1.1	Accreditation	96
4.2.1.2	Quality systems	97
4.2.1.3	CHI and CHAI	97
4.2.1.4	Performance management: the PAF	99
4.2.1.5	Performance ratings ('stars')	100
4.2.1.6	National Service Frameworks (NSFs)	101
4.2.2	External / formative	101
4.2.2.1	Quality management (EFQM)	101
4.2.2.2	Benchmarking processes	103
4.2.2.3	Professional perspective (<i>Visitatie</i>)	104
4.2.2.4	Modernisation Agency (MA)	105
4.3	Internal approaches	105
4.3.1	Internal / summative	105
4.3.2	Internal / formative	106
4.3.2.1	Statistical process control	106

		Page
4.3.2.2	Specific clinical governance measures:	107
4.3.2.2.1	Clinical governance self-assessment and development guides (WMRO)	107
4.3.2.2.2	CHI Self-assessment schedules	108
4.4	The need for further measures	109
4.4.1	The case for a measure of clinical governance climate	109
4.4.2	The case for a measure of organisational competency	119
	Conclusions	123
Chapter 5	Measures of clinical governance I: The Clinical Governance Climate Questionnaire (CGCQ)	124
	Introduction	
5.1	An overview of CGCQ development methods	125
5.2	Literature review to identify potential measurement domains	129
5.3	Qualitative development of the CGCQ item pool	132
5.3.1	Focus groups	133
5.3.2	Methodology	143
5.3.3	Results	153
5.4	Draft indicator item pool based on focus group results	183
5.5	Content validation of the CGCQ items	185
5.5.1	Content validation	185
5.5.2	Methodology	188
5.5.3	Results	191
5.6	Item pool reduction	197
5.6.1	Application of exploratory factor analysis	198
5.6.2	Methodology	219
5.6.3	Results	229
5.7	Internal consistency (reliability) and external validity	240

		Page
5.7.1	Reliability and validity	240
5.7.2	Internal consistency (Cronbach's Alpha)	248
5.7.3	External validity	250
5.7.3.1	Discriminant validity	250
5.7.3.2	Construct validity	254
5.8	Discussion	262
	Conclusions	265
Chapter 6	Applying the CGCQ: inter-organisational and inter-professional effects	266
	Introduction	
6.1	Methodology	267
6.2	Results	270
6.2.1	Descriptive detail on study sites	270
6.2.2	Analytical results	273
6.3	Discussion	279
	Conclusions	286
Chapter 7	Measures of clinical governance II: the Organisational Progress in Clinical Governance (OPCG) schedule	287
	Introduction	
7.1	Developing a competency framework	287
7.1.1	An overview of OPCG development methods	288
7.1.2	Operationalising good practice: generating the tasks	292
7.1.3	Operationalising levels of attainment	298
7.1.3.1	Use of Q-sort in the present study	299
7.1.3.2	Q-sort study methodology	304
7.1.3.3	Results	310

		Page
7.2	Reliability and validity of the OPCG Schedule	314
7.2.1	Reliability (internal consistency): factor analysis and co-efficient alpha	314
7.2.1.1	Factor analysis	314
7.2.1.2	Co-efficient alpha	318
7.2.2	Validity	318
7.2.2.1	Content validation	318
7.2.2.2	Construct validation	322
	Conclusions	323
Chapter 8	Applying the OPCG competency framework: a national study of progress in clinical governance	325
	Introduction	
8.1	Methodology	326
8.2	Results	330
8.3	Discussion	380
	Conclusions	390
Chapter 9	Conclusions	391
	Introduction	
9.1	Progress in clinical governance	392
9.2	Options for evaluating the effects of CHI reviews	395
Appendix 1	A digest of empirical studies of clinical governance implementation in primary and secondary care	408
Appendix 2	A digest of primary and secondary material on the use of performance indicators to improve health care quality	415
	2a: Use of performance indicators in improving health care quality (general)	416

	Page
2b: Use of performance indicators in improving health care quality (outcome indicators)	425
Appendix 3 Raw data from the focus groups: abridged data under domain 1 (adverse events)	434
Appendix 4 The CGCQ validation study questionnaire	460
Appendix 5 CGCQ validation study tables	491
Appendix 6 The CGCQ schedule used in the development study sample	521
Appendix 7 The 60-item short form CGCQ	538
Appendix 8 Skewness and kurtosis of CGCQ items from the development sample pre- and post-transformation	542
Appendix 9 Complete rotated factor scores for six-factor PFA with Varimax rotation	547
Appendix 10 Instructions to Q-sort participants	550
Appendix 11 Q-sort results	555
Appendix 12 The OPCG schedule	566
Appendix 13 OPCG items by structure, process or outcome	574
Appendix 14 Step-down F analyses of trust type, CHI visit and board membership, using OPCG domain scores as independent variables	575
Appendix 15 Publications derived from the thesis	576
Bibliography	577

	List of Illustrations	Page
Figure		
1.1.2.1	Modes of governance adapted and extended from Hood <i>et. al.</i> (1999)	10
1.1.4.1	Meanings of modernisation, in Peck (2000)	19
1.1.4.2	The UK health care quality strategy, as outlined in 'A first class service: quality in the new NHS' (DoH 1998)	20
1.1.4.3	The UK health care quality strategy after the NHS Plan, 2000	23
4.1.1	Control locations and resultant action matrix (Boland and Fowler 2000)	90
4.1.2	Control locations and resultant action matrix in the UK NHS	94
4.2.2.1.1	The EFQM Excellence Model	102
4.4.1.1	Conceptual distinctions between 'culture' and 'climate' research	115
4.4.1.2	Quantitative organisational culture measures, after Scott <i>et. al.</i>	117
5.6.2.1	Bivariate barcharts of CGCQ item distributions	228
5.6.3.1	Scree-plot of factors against eigenvalues	232
5.7.3.1.1	Inter-trust comparison: mean factor scores	252
5.7.3.1.2	Profession-specific assessment of climate: mean factor scores	253
6.2.1.1	CVQ culture sub-scale scores by trust	273
7.2.1.1	Scree-plot of factors against eigenvalues	316
7.2.2.2.1	Construct validation of levels of difficulty: Mean achievement scores for items indicative of low, medium and high achievement	323
8.2.1	Perceived importance of competencies	331
8.2.2	Achievement in competencies	333
8.2.3	Shortfalls in competencies	335
8.2.4	Quality improvement items: importance / achievement / shortfall	338
8.2.5	Risk management items: importance / achievement / shortfall	341
8.2.6	Performance improvement items: importance / achievement / shortfall	344
8.2.7	Corporate accountability items: importance / achievement / shortfall	347
8.2.8	Leadership & collaboration items: achievement / importance / shortfall	350
8.2.9	Perceived achievement, importance and shortfall across domains	353
8.2.10	Perceived achievement against structure, process and outcome	354
8.2.11	Ranked aggregate domain scores by trust	356
8.2.12	Interaction effects between trust type and CHI visit-status on improving performance	366

	List of Illustrations	Page
Figure		
8.2.13	Interaction effects between trust type and CHI visit-status on managing risks	367
8.2.14	Marginal means of structure	373
8.2.15	Marginal means of process	374
8.2.16	Marginal means of outcome	375
Box		
4.4.2.1	Data collection approaches	121
5.6.1.1	Total variance: equation 1	205
5.6.1.2	Total variance: equation 2 (after Child, 1976)	205
5.7.1	The relationship between latent variable, indicator items, and unique variance (error)	242

	List of Tables	Page
1.1.1.1	Clinical governance: definitions	5
1.1.5.1	Second generation NHS regulatory mechanisms	27
3.3.1.1	Accountability and improvement approaches to indicator systems	71
3.3.2.1	Unintended consequences of public sector performance indicator systems	80
4.4.1.1	Conceptual distinctions between ‘culture’ and ‘climate’ research	115
4.4.1.2	Quantitative organisational culture measures	117
5.1.1	CGCQ development research programme	126
5.2.1	Potential measurement domains for clinical governance climate	131
5.3.1.1	Focus group definitions	134
5.3.1.2	Methodological tensions within normative focus group literature	138
5.3.2.1	Focus group topic guide	145
5.3.2.2	Composition of focus groups, by site	148
5.3.2.3	Number of data segments, by domain	152
5.3.3.1	Domains and their inductively generated sub-themes	154
5.3.3.1.1	Openness and encouragement to report adverse events	155
5.3.3.1.2	Blame	156
5.3.3.1.3	Using a broad approach in investigation	157
5.3.3.1.4	Making changes where necessary	158
5.3.3.2.1	System-wide orientation	159
5.3.3.2.2	Stakeholder involvement	160
5.3.3.2.3	Multi-disciplinary focus	161
5.3.3.2.4	Investment in improvement	162
5.3.3.2.5	Innovatory environment	163
5.3.3.2.6	Use of quality improvement techniques	164
5.3.3.2.7	Planned improvement	165
5.3.3.3.1	Assessment	166
5.3.3.3.2	A corporate approach	167
5.3.3.3.3	Action on identified risks	168
5.3.3.4.1	Investment in learning	169
5.3.3.4.2	Sharing learning with others	171
5.5.5.5.1	Motivation to change	172
5.3.3.5.2	Support	173
5.3.3.5.3	Leadership	174
6.3.3.6.1	Technical skills	175
5.3.3.6.2	Access to research evidence	176
5.3.3.7.1	Common goals	177
5.3.3.7.2	Mutual respect	178
5.3.3.7.3	Clear roles	179
5.3.3.7.4	Participation	180
5.3.3.8.1	Reviews used developmentally	181
5.3.3.9.1	Training and development needs	182

	List of Tables	Page
5.3.3.9.2	Using new skill	183
5.4.1	Characteristics of good indicators based on DeVellis (1991) and Oppenheim (1992)	184
5.5.3.1	Typology of items identified by respondents as cause for concern	192
5.5.3.2	Typology of ambiguities in the original item pool	194
5.5.3.3	New items suggested by respondents included in item pool	195
5.6.1.1	Factor analysis definitions	200
5.6.2.1	Data screening (after Tabachnick and Fidell 2001)	223
5.6.3.1	Respondent characteristics	230
5.6.3.2	Total Variance Explained, initial 20 factor	232
5.6.3.3	Factor model adequacy	234
5.6.3.4	Factor Correlation Matrix	235
5.6.3.5	Percentage of variance and covariance explained by rotated factors	236
5.6.3.6	Factor loadings and communalities (h^2) for Principal Factors extraction and Varimax rotation of CGCQ (factor loadings < .42 omitted)	238
5.7.2.1	Cronbach's alpha coefficients for CGCQ sub-scales	248
5.7.3.1.1	Discriminant validity: differences and similarities in CGCQ factor scores by profession and trust	251
5.7.3.2.1	Intra-scale correlation: CVQ culture-type and MBI sub-scales (Pearson's r)	257
5.7.3.2.2	Concurrent validity: correlation of CGCQ with CVQ and MBI sub-scales (Pearson's r)	261
6.2.2.1	Multivariate tests of significance (Pillai's trace and Wilkes' lambda)	274
6.2.2.2	Correlations between dependent variables (Pearson's r)	275
6.2.2.3	Univariate F-scores of effect of trust on CGCQ factor scores	275
6.2.2.4	Inter-trust means for organisational learning, training and development and quality improvement	276
6.2.2.5	Univariate F-scores of effect of professional group on CGCQ factors	277
6.2.2.6	Inter-professional means for absence of blame, risk management and working with colleagues	278
7.1.1.1	OPCG competency framework research programme	289
7.1.2.1	OPCG Task areas mapped on to 'Main components of clinical governance for NHS Trusts' from discussion and guidance documents, DoH 1998/9	295
7.1.2.2	Benefits arising from clinical governance- the vision for the next five years (after <i>Clinical governance: quality in the new NHS</i>)	297
7.1.2.3	The 20 OPCG tasks, derived from discussion and guidance documents	298
7.1.3.2.1	Q-set generation study participants	306

	List of Tables	Page
7.1.3.3.1	Q-sort items retained	312
7.1.3.3.2	Validated competence level	313
7.2.1.1.1	Total variance explained, initial 5 factors, pre- and post-rotation	316
7.2.1.1.2	Factor loading and communalities (h^2) for principal factors with varimax rotation of OPCG items	320
7.2.1.2.1	Alpha co-efficients of the OPCG domains	318
8.2.1	Perceived importance of competencies	332
8.2.2	Achievement in competencies	334
8.2.3	Shortfall in competencies	336
8.2.4	Quality improvement achievement data	339
8.2.5	Risk management achievement data	342
8.2.6	Performance improvement achievement data	345
8.2.7	Corporate accountability achievement data	348
8.2.8	Leadership & collaboration achievement data	351
8.2.9	Analysis of variance: achievement scores by trust (1-100)	355
8.2.10	Multivariate tests of significance (Pillai's trace and Wilkes' lambda)	359
8.2.11	Correlations between dependent variables (Pearson's r)	360
8.2.12	Tests of trust type, CHI visit and board membership, and interactions (OPCG domains)	362
8.2.13	Between trust variation	363
8.2.14	Between CHI visit-status difference	364
8.2.15	Between board-member status differences	364
8.2.16	Interaction effects	365
8.2.17	Multivariate tests of significance (Pillai's trace and Wilkes' lambda)	368
8.2.18	Tests of trust type, CHI visit and board membership, and their interactions (structure / process / outcome aggregates)	369
8.2.19	Correlations between dependent variables (Pearson's r)	368
8.2.20	Trust differences	370
8.2.21	CHI visit-status differences	370
8.2.22	Boardmember status differences	371
8.2.23	Estimated marginal means	372
8.2.24	Multivariate tests of significance (Pillai's trace and Wilkes' lambda)	376
8.2.25	Correlations between dependent variables (Pearson's r)	377
8.2.26	Tests of trust type, CHI visit and board membership (low / medium / high aggregates)	377
8.2.27	Trust differences	378
8.2.28	CHI visit-status differences	379
9.2.1	Five principal axes of evaluation research	396
9.2.2	Five principal axes of evaluation, by 'ideal-type' paradigm	402

CHAPTER 1

CLINICAL GOVERNANCE: A CRITICAL INTRODUCTION

"... a history respectful of contingency, mistrustful of inevitability, indifferent to any predetermined route or destination; a history refusing to take for granted (as the victors' texts always want) that the way things turned out was the way they were always meant to be..." (Schama 2000: p. 17)

"A consideration of clinical governance involves engaging with issues of autonomy and control in the context of the complex relationships that exist between clinicians and managers and between policy makers and service providers. It also involves unpacking the vague, or contested, idea of governance itself..." (Walsh & Small 2001: p. 110)

"It includes all of the favoured approaches to performance measurement (the approach that one of my colleagues referred to as 'hitting the donkey with the carrot')..." (Broadbent 2003: p. 5)

Introduction

'Official' accounts present clinical governance as an inevitable government response to wide variations in clinical practice and the loss of public confidence occasioned by a series of scandals over health care quality (Sally & Donaldson 1998; Nicholls *et. al.* 2000).

While such high profile scandals undoubtedly made health care quality a political imperative and precipitated a response, the official view does little to explain the *specific form* of the prescription. In order to better explore continuities and discontinuities with previous governance arrangements in the UK NHS and the difficulties posed by its internal tensions, clinical governance is situated in the context of broader shifts in modes of governance between the state, citizens and professional groups. This chapter locates

clinical governance within the broader set of health care quality reforms contained in the 1997 White Paper *The new NHS: modern, dependable*, providing a critical overview of the concept as outlined and further developed in government discussion and guidance documents (DoH 1998; NHSE 1999).

The exponential growth in external regulatory mechanisms throughout western Europe since the 1990s has been occasioned by increased decentralisation and use of regulated markets in service delivery, with consequential greater emphasis on oversight methods of accountability (Power 1997; Hood *et. al.* 1999). Debates over models and methods of external review are ultimately concerned with professional control, and the perceived value of approaches is dependent upon judgements concerning the appropriate balance between professional self-regulation and accountability as their use strikes new balances of trust and control between patients, professionals and government.

Within the UK the role of central government in health care quality improvement has become increasingly prescriptive. Six years have passed since the government outlined its programme of NHS quality reforms (DoH 1997), an ambitious agenda reflecting broader governance trends towards the performance management of devolved organisations against national standards (Newman 2000; Geddes & Martin 2000; Ball *et. al.* 2002; Freeman 2002; Broadbent 2003). New NHS systems outlined in the 1997 White Paper and later NHS Plan (2000) included a National Institute for Clinical Excellence (NICE) charged with standard setting; an expanded Performance Assessment Framework; visits, inspections and latterly performance ratings ('stars') from a Commission for Health care Audit and

Inspection (CHAI); and a Modernisation Agency to support organisational process redesign.

Underpinned by a new statutory duty for quality, structures of clinical governance within health care provider units are designed to ensure local delivery of the national agenda. The lynchpin of the reforms, clinical governance is defined as:

"... a framework through which NHS organisations are accountable for continuously improving the quality of their services and safeguarding high standards of care by creating an environment in which excellence in clinical care will flourish"
(DoH 1998: p.33).

Given ambiguities in the definition, the term clinical governance has been strongly contested and subject to multiple interpretations ever since it was first coined. It has been characterised as both a controls assurance framework (Emsley 2002; Martin 2001) and a whole-systems approach to continuous quality improvement (Latham *et. al.* 2000; Halligan and Donaldson 2001; Wallace *et. al.* 2001; Cullen *et. al.* 2000; Hackett *et. al.* 1999; Nicholls *et. al.* 2000).

How are such divergent readings possible? Support for both may be found within the White Paper (DoH 1997). While short-term targets relate to development of assurance committee structures, the long-term vision is fixed on cultural change, away from a climate of blame and fear towards open reporting of incidents, reflective learning and service development. This is reinforced by the five-year vision for clinical governance outlined in the guidance document (NHSE 1999), supporting the development of open, participative

cultures, active user involvement, multidisciplinary team-working and quality improvement. Yet, the language of continuous quality improvement contrasts sharply with the assurance-focused style of performance management exhibited by the Department of Health and many of its agencies, and there is a danger that the long-term quality improvement agenda is being diluted.

1.1 Clinical governance, its origins and status

1.1.1 Definitions

Since its emergence on publication of the 1997 White Paper '*The new NHS: modern, dependable*', clinical governance has demanded the attention of managers, clinicians and academics alike. Published responses cover the spectrum from rejection (Goodman 1998) via cautious approval (Davies & Mannion 1999; Davies *et. al.* 2001, Walshe *et. al.* 2000), to enthusiastic endorsement (Donaldson & Muir-Gray 1998; Scally & Donaldson 1998). Such divergences of opinion are due to contested understandings of the 'nature of the beast'.

The contrast between the definition provided by DoH 1997 and DoH 2002b on one hand and NHSE 1999 on the other (table 1.1.1.1) encapsulate the tensions between competing claims made on behalf of clinical governance. While the former definitions locate the essence of clinical governance within local structures and processes backed by a statutory duty, the latter locates it within a change in organisational culture towards greater openness, learning and service improvement.

Table 1.1.1.1: Clinical governance: definitions

DoH (1997:3.6)	"A new system of clinical governance in NHS Trusts and Primary care to ensure that clinical standards are met, and that processes are in place to ensure continuous improvement, backed by a new statutory duty for quality in the NHS."
DoH (1998: p. 33)	"... a framework through which all National Health Service (NHS) organisations are accountable for continuously improving the quality of services and safeguarding high standards of care by creating an environment in which excellence in clinical care will flourish."
NHSE (1999: p.5)	"Above all clinical governance is about changing organisational culture in a systematic and demonstrable way, moving away from a culture of blame to one of learning so that quality infuses all aspects of the organisation's work."
Scally & Donaldson (1998: p. 62)	"Clinical governance is a system through which NHS organisations are accountable for continuously improving the quality of their services and safeguarding high standards of care by creating an environment in which excellence in clinical care will flourish."
Goodman (1998: p. 1726-7)	"... a rolling unpunctuated tangle of prepositional and adverbial phrases similar to the many that have appeared in the past few years on the walls of our hospital wards and clinics ... a small idea, a rehash of all sorts of "management speak" and poorly thought through generalisations that can depress and dishearten."
Onions (2000: p. 405)	"... a new NHS concept that seeks, through clinical guidelines and national service frameworks, to encourage and enforce compliance with nationally-devised evidence-based clinical policies."
DoH (2002b: p. 20)	"... the local delivery mechanism for ensuring safe and high quality care. Clinical governance will ensure that NHS organisations have in place systems to assure the quality of their services and processes to ensure continuous improvement year-on-year."

Indeed, close reading of the definition of clinical governance within the discussion document (DoH 1998: p.33) offers further insight into its contested nature. The definition

contains elements of both external quality assurance and internal continuous quality improvement, which offer significantly different visions of the best approach to health care quality. To the extent that clinical governance provides an accountability framework of standards against which organisations and / or individuals may be assessed by external agencies, it constitutes a performance management system. On the other hand, to the extent that it relies on development and support to encourage continuous quality improvement it constitutes a whole-systems approach to Continuous Quality Improvement (CQI).

While both assurance and improvement approaches rely on indicators, they use them to different ends; the former to make summative judgements on care quality, the latter to foster learning and improvement within clinical teams (Freeman 2001). The dual emphasis on assurance and improvement found within clinical governance has been described as bridging the gap between managerial and clinical approaches to quality (Buetow & Roland 1999). However, the difficulty is that there is no attempt to dialectically resolve the tensions between the two:

“The performance focus contrasts sharply with the language of continuous quality improvement, and the interplay of these two discourses informs activity at the sharp end of service delivery.” (Newman 2000: p.55)

The resulting struggle for the soul of clinical governance is summarised in a series of essays entitled 'clinical governance: enabling excellence or imposing control?' (Miles *et. al.* 2000). While quality improvement and central control are conceptually distinct and are not strictly antonyms (Harrison & Dowswell 2000), the fear is that performance management

approaches may undermine the conditions of trust, honesty and innovation required for quality improvement (Davies & Mannion 1999a). The balance between checking and trusting becomes critical, as an overemphasis on performance management risks the long-term quality improvement agenda.

For Walshe (2000), the inclusion of previous policy initiatives such as risk management, clinical audit and evidence-based practice in descriptions of clinical governance within government documents belies its newness. Despite its lack of methodological innovation, clinical governance is not simply an exercise in 're-badging' or re-branding. The integrationist description misses the central point of *governance*: while comprising many of the same elements as previous initiatives, clinical governance imposes a statutory responsibility for quality on individuals and organisations. For the first time, individuals and organisations are accountable for achieving central quality targets (Ferguson & Lim 2001). These requirements are also explicitly linked to performance assessment mechanisms:

"Although the emphasis on controlling, guiding, influencing, managing or regulating may differ according to who is doing the "governing", the underlying rationale is clear: individuals and / or organisations need to be made more accountable in order to achieve outcomes that are deemed desirable."
(Ferguson & Lim 2001: p. 464).

The study of clinical governance requires exploration of the concept of governance, and of broader shifts in governance modes, as well as its position within the 1997 White Paper reforms. Clinical governance is best understood in the context of a range of new external

institutions and relationships, constituting a strengthening of performance management instruments for UK health care. These new institutions were explicitly designed to ensure optimal performance of the NHS in line with UK policy directives, implying both the primacy of managerial as opposed to clinical or patient actions and involvement, and the 'given' status of policy objectives (Smith 2002).

Such an examination of clinical governance in the context of wider governance trends (Newman 2000, Power 2000) offers a powerful revisionist critique of the official history of clinical governance as an inevitable response to failures in care quality, wide practice variations and falling public confidence (Nicholls *et. al.* 2000). The approach locates clinical governance within a broader historical and conceptual context to explain its specific form, internal tensions and similarities with parallel developments in other policy arenas (Geddes & Martin 2000; Springings 2002).

1.1.2 Governance and modes of governance

Rhodes (1996) identifies two broad strands of definitions of 'governance'. The first is associated with reducing the role of the state and employing new public management techniques in order to do more with less (Osborne & Gaebler 1992), and the second with theories of the interdependence of public, private and semi-private actors, such as self-organising networks. Both approaches are sceptical of the ability of a central state to directly plan economic and social organisation, and shift the emphasis from government

(noun) to governance (verb). However, they differ fundamentally in their perspective on public management and the role of government in society. In short, while the former is an attempt to translate managerial ideas such as contracting out and client orientation from the private sector to public services, the latter focuses on co-ordinating inter-organisational policy-making (Klijn & Koppenjan 2000). While partnership working between health and social care agencies has been an important part of New Labour's governance strategy, the emphasis has been on an inclusive managerialism rather than community activism (Walsh & Small 2001).

For Kooiman (1993), governance is the emerging pattern that arises from governing- the purposive means of guiding and steering a community, organisation or society.

Governance then is a form of social co-ordination (Mayntz 1993), of which many patterns, or modes, are available. Lowndes & Skelcher (1998) acknowledge the intellectual debt owed to Williamson's analysis of markets and hierarchies as governance structures (Williamson 1985) in all attempts at differentiation of modes of governance. They identify a number of triads from the literature, including markets, hierarchies and networks (Thompson *et. al.* 1991); markets, bureaucracies and clans (Ouchi 1979; 1991); and community, market and state (Streek & Schmitter 1985). Hood *et. al.* (1999) propose four: oversight (bureaucracy), mutuality (clan), competition (market), and contrived randomness. While there are differences of emphasis in these analyses, four ideal types emerge: competition, oversight, network and mutuality (figure 1.1.2.1).

Figure 1.1.2.1: Modes of governance, adapted and extended from Hood *et. al.* (1999)

Competition [market] control through rivalry and choice	Oversight [bureaucracy, hierarchy] command and control
Mutuality [professionalism] control through group processes	Network [interdependency] control through mutual benefits

Competition / market modes of governance

Market modes of governance deploy contractual relationships between parties, mediated through price mechanisms. While they provide flexibility to actors in terms of their willingness to form alliances, the competitive environment means that parties will generally only collaborate where there are clear mutual benefits, given underlying suspicion between parties. Co-ordination and collaboration are thus difficult. In the NHS, this mode of governance is most closely associated with the internal market reforms of the early 1990s. The distinction between purchasers and providers, in which providers competed for state finance, was intended to produce market incentives to contain costs and increase service quality. While competition in the NHS was severely constrained by a number of factors, including the geographical proximity of hospital providers to purchaser populations and political interference, the purchaser / provider split was retained in the New Labour reforms. Indeed, *Shifting the Balance of Power* further extends to purchasing power of Primary Care Trusts (PCT's) to purchase care from the secondary acute sector.

Oversight / bureaucracy / hierarchical modes of governance

Oversight modes of governance rely on an authoritative supervisory structure to overcome the difficulties of co-ordination and collaboration found in the market. However, regulation comes at the cost of stifling flexibility and innovation, given the formality and routine nature of hierarchical relationships. Oversight relies on rules and targets enforced by systems of hierarchical surveillance- each subordinate is assessed against standards set by their superior in a hierarchy. Consequently, performance is assessed in terms of compliance with rules and meeting targets. While oversight and has been used extensively in the NHS, historically the focus has been on activity and finance, with clinical quality largely left to clinical professionals (Davies & Mannion 1999b).

Mutuality / clan / professional modes of governance

Mutuality or clan control relies on trusting group norms or professional values to control individual behaviour- individuals are thus controlled through group influences, via the conscious pursuit of reputation among peers over time. Such methods of control are seen as particularly effective in circumstances where individuals are bound by strong professional norms, success is difficult to measure, and where the relationship between inputs, processes and outcomes is unclear (Ouchi 1979). Indeed, neither self-interested contracting (market) nor hierarchical enforcement can be relied upon when it is not possible to specify in detail what constitutes as 'good job'. In such circumstances, a horizontal organisation, or professional clan, needs to be involved. Thus health care relies on validation by clinical professions, especially medicine. Professions are themselves

organisational forms, and mutuality dominated the approach to the control of clinical quality within the NHS until the 1997 reforms. Indeed, the clinical autonomy of the medical profession to assess performance via peer review is a founding principle of the NHS.

Network governance

Based on the 'governing without government' thesis of Rhodes (1996; 1997; 1999), network governance contends that governing increasingly depends on interactions between public and private sector actors which are relatively autonomous of central states. Consequently, central states can steer only indirectly and imperfectly. Actors are seen as capable of identifying complementary interests, and developing interdependent partnerships based on trust, reciprocity and a sense of common purpose (Lowndes & Skelcher 1998). While partnerships have themselves been characterised as a mode of governance (Davies 2002), Lowndes & Skelcher (1998) contend that partnership arrangements draw on hierarchy, network and market modes of governance at different stages of development.

The typologies above are best viewed as Weberian 'ideal types'; in practice they are not mutually exclusive and tend to function in parallel as hybrids. Hood *et. al.* (1999) cite the rating of universities for teaching and research quality by assessment panels made up of assessors drawn from teaching and research as an example of an oversight / mutuality governance hybrid. In the context of clinical governance, the inclusion of clinical, managerial and patient representatives on Commission for Health Improvement review teams could be viewed as a oversight / mutuality hybrid. The use of league tables to assess

provider institutions, as in the Performance Assessment Framework, could be seen as an example of an oversight / competition hybrid. On this analysis, clinical governance is not itself a mode of governance- a means by which social co-ordination is achieved. It is less a new 'post-bureaucratic' form (Hoggett 1996) than a shifting blend of different forms of social co-ordination, including market, hierarchy, clan and network modes of governance. Indeed, the absence of detailed prescription in the form of clinical governance arrangements in the discussion and guidance documents provided some latitude for local interpretation in implementation, which is explored in chapter two.

1.1.3 Principal – agent theory

The dilemmas facing those charged with influencing the behaviour of autonomous professional groups may be usefully explored from the economic perspective of principal-agent theory (Davies & Mannion 1999a; 2000). According to this perspective, principals are those individuals or agencies who want things done and agents are those who are engaged in accomplishing tasks for the principals. While such relationships may be contractual or informal, a key feature of all principal-agent relationships is the presence of asymmetries between the parties, leading to the potential for agent opportunism.

From this perspective, clinicians are in multiple principal-agent relationships with patients, managers, and commissioners (Smith *et. al.* 1997; Shortell *et. al.* 1998). In the context of clinical governance it is helpful to consider health service managers with responsibility for

health care quality as principals, and the clinicians through whom care quality may be either achieved or lacking as the agents of managers. Davies & Mannion (2000) identify two important differences between such principals and agents: objective functions, and knowledge and information. The former refers to ambiguity over the objectives of health care that may lead to divergence between managers and clinicians. A good example here is the primacy of balanced budgets for health care managers, whereas clinicians may be seeking to maximise health benefits for the patients in their care. The latter refers to the different technical knowledge and situation-specific information available to principals and agents, which may lead to differences in understandings and judgements of appropriate action. Such differences provide scope for divergence between what principals want and what agents deliver:

“Agents may for example make undesirable trade-offs between multiple competing objectives, or worse, they may exploit their powerful position to indulge in opportunistic behaviour to maximise their own gain at the expense of the principal.”
(Davies & Mannion 2000: p. 252)

While the opportunistic behaviour of professional agents is likely to be curtailed by self-community- and client-control (Sharma 1997), two major approaches are available to principals to reduce the potential for dysfunctional agent behaviour. The first consists of the checking and modification of behaviour using hierarchical control structures in an attempt to redress asymmetries of information. This requires the accurate measurement of the aspects of agent behaviour deemed important, and ability to exert influence over agent behaviour that does not conform to the measurement ideals. In contrast, the second

approach is formative, seeking the negotiated realignment of objectives and beliefs that might give rise to agent waywardness. This approach seeks to develop agents' intrinsic professional motivations through the development of trust, in order to realign agent objectives with those of principals.

Managers charged with service improvement are likely to be drawn instinctively to coercive methods of behaviour change. Yet, a focus on data to drive improvement suffers from difficulties of collection, interpretation, and perverse incentives, which are explored in detail in chapter three. Given the difficulty of measuring clinical outcomes and difficulties in attribution between clinician behaviour and clinical outcomes, Davies & Mannion (1999a) caution against the over-reliance on oversight to redress information asymmetry between principals and agents:

“The concern is that insufficient attention will be paid to the sometimes nebulous concepts of trust and culture in a headlong rush for the more tangible appeals of measurement, monitoring and coercive control mechanisms.”
(Davies & Mannion 1999a: p.15)

1.1.4 Historical development of modes of governance in the UK NHS

While the number of ideal types of governance modes is limited, the relative balance between these approaches and their hybrids is much more variable. Since the UK NHS's inception, the balance between different forms of governance has shifted over time reflecting broader governance trends. Three periods may be distinguished: the 1948

settlement (command and control and clan); New Public Management under the 1991 reforms (market and oversight); and New Labour's 'third way' set out in the 1997 White Paper (eclecticism).

(a) the 1948 Settlement

At its inception, the NHS relied on a blend of hierarchy and clan modes of governance. For major decisions about priorities and provision the NHS relied on an underlying hierarchical system of command and control, with strong line management and upward accountability (Robinson 2002). Notwithstanding the latitude for devolved decision-making at regional, area and unit level and the introduction of general management in 1984, the predominant mode of governance was hierarchical. Command and control was tempered by clan modes of governance by the medical profession (Bourne & Ezzamel 1986), in which the professional training and specialised knowledge of clinicians was perceived to make them best placed to judge clinical performance and the quality of clinical care (Goddard *et. al.* 2000).

(b) New Public Management (NPM) under the 1991 reforms

In the 1980s, New Right political theory questioned state public services' efficiency and service quality. Public bureaucracies were characterised as unable to address the concerns of service users, which would be better served through the application of 'entrepreneurial governance' forms (Osborne & Gaebeler 1992). This approach conceptualised clients as customers, and advocated decentralisation to competing, autonomous local providers regulated by market mechanisms. The application of these ideas to UK NHS health

services resulted in the introduction of the internal market in 1991. The main feature of the approach was to separate the responsibility for purchasing services from their provision, by charging fund holding GPs and Health Authorities with the responsibility of commissioning services and creating autonomous health care trusts (providers). Potential service providers were then encouraged to compete against each other to win the contract for provision. This devolved decision-making to purchasers and providers, so that accountability was essentially downward to service consumers. While the system was intended to introduce market forces into public provision, the demands of political accountability quickly led to close regulation and management of the health care internal market via systems of inspection, accounting, regulation and review (Clarke *et. al.* 2000).

The rise of semi-autonomous, decentralised provider units subject to 'hands-off' central performance management is explored in the social policy literature under the title of New Public Management (NPM) and the regulatory state. The term 'regulatory state' suggests the increasing use of standards, rules and monitoring, using public authority to apply rules and standards. Hood *et. al.* (1999) offer an account of the changes in public sector regulation during the 1980s and 1990s, and contends that regulation increased in formality, complexity intensity and specialisation throughout the period, due to the rise of New Public Management:

“New Public Management stressed the importance of [performance accountability], emphasising a change from tactical to strategic prescription, direct to indirect command, detailed instruction to freedom within constraints. Even on that characterisation, new regimes of performance regulation can be expected to emerge in parallel with new management freedoms, and ‘compliance’ regulation [over procedures, for instance on

merit-hiring or complaint handling] *can take new, often enhanced, forms in a more managerial service.*"

(Hood *et. al.* 1999: p. 6)

Hoggett (1996) identifies three defining interlocking strategies of NPM. The first consists of the drive towards operationally decentralised units, with a simultaneous drive for increased central control over strategy and policy. Secondly, the activities of decentralised units were co-ordinated via competition and/ or market incentives. Finally, the activity of the decentralised units increasingly became subject to performance management and monitoring via audits inspections and reviews. The simultaneous use of both centralisation and decentralisation are a central paradox of the approach (Clarke and Newman 1997), and the need for regulatory mechanisms arises due to the need to 'steer' the behaviour of semi-autonomous organisations. Consequently, NPM is best seen as a combination of competition and oversight modes of governance, the former via contracting and the latter via systems of performance management:

"The picture that therefore emerges is one of an operationally decentralised organisation with a strong but distant centre engaging in performance monitoring and shaping activity by concentrating on a few indicators which give emphasis to results rather than inputs or processes."

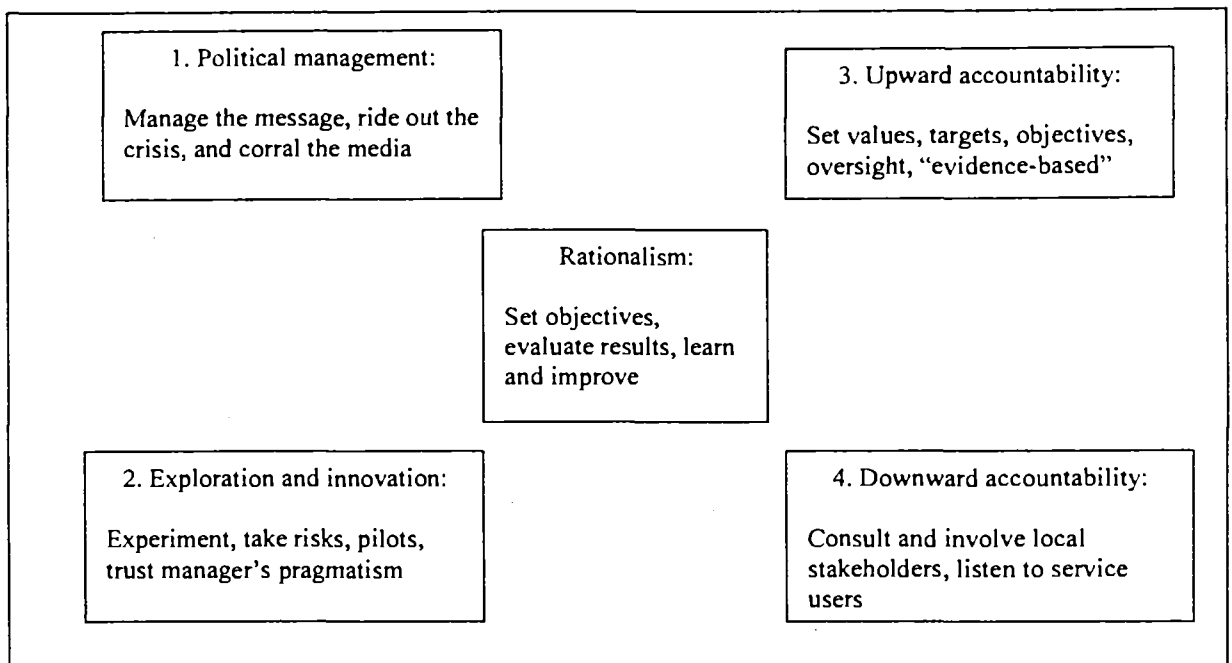
(Hoggett 1996: p.21)

(c) 'Modernisation'- New Labour's third way

At the heart of the modernisation project lie tensions between central prescription, local innovation and upward accountability (Peck 2001; 6 2000). 6's analysis is particularly

helpful (figure 1.1.4.1) in drawing a distinction between local exploration and innovation sanctioned by the centre (item 2), and a pluralist downward accountability routinely seeking local innovation (item 4). The former is consistent with earned autonomy under a performance management framework, the latter with network modes of governance.

Figure 1.1.4.1: Meanings of modernisation, in Peck (2000)

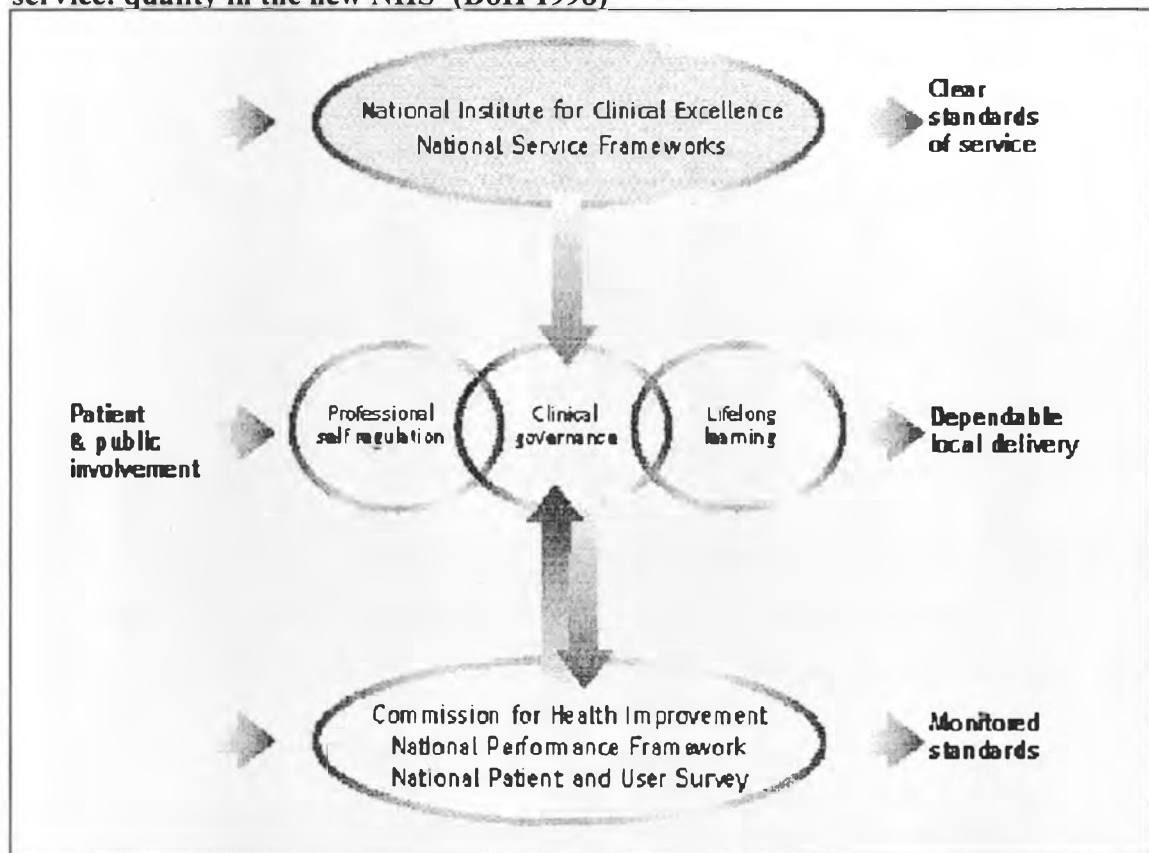


Two distinct phases of New Labour's third way may be discerned, suggestive of an emergent strategy seeking to balance the above tensions, rather than an ideological commitment. The first phase is codified in the 1997 White Paper *The new NHS: modern, dependable* (DoH 1997), and the second in the subsequent *NHS Plan* (DoH 2000).

(i) Phase I: The New NHS: modern, dependable 1997 ('Modern Public Management')

New Labour vehemently opposed the internal market while in opposition, and instead offered a vision of partnership working and collaboration between health and social care providers. However, New Labour's third way retained and extended the emphasis on decentralisation and performance management, with additional elements of central direction (Figure 1.1.4.2). Commentators have satirised its continuities with New Public Management by coining the term 'Modern Public Management' (Newman 2000).

Figure 1.1.4.2: The UK health care quality strategy, as outlined in 'A first class service: quality in the new NHS' (DoH 1998)



Under the 1997 reforms care standards were set nationally via a new National Institute for Clinical Excellence (NICE) and new National Service Frameworks (NSFs). National standards were monitored through performance management and external regulation via a rolling programme of visits from a new Commission for Health Improvement (CHI) and a series of high level performance indicators under a Performance Assessment Framework (PAF). These new developments are considered in further detail in section 1.1.3 below and in the context of measuring progress in clinical governance in chapter 4. The role of clinical governance within the reforms was to ensure the local delivery of the national agenda, through internal committee systems and by engaging clinical professionals in quality improvement activities; a dual strategy of internal quality improvement and external quality assurance.

While self-regulation remains in name (Figure 1.1.4.2), it is 'modernised' to be open to scrutiny and publicly accountable for nationally set professional standards (DoH 1998: para 3.44). In other words, the scope for oversight modes of governance in connection with clinical quality has been dramatically increased. Oversight and clan modes of governance in the context of professional standards make potentially uneasy bedfellows:

".... if necessary clinical governance will provide a systematic framework that can be extended into the clinical community at all levels. Here we have the rationalistic bureaucratic discourse of regulation, which reveals itself through increasingly extensive rule systems, the 'scientific' measurement of objective standards, and the minimisation of the scope for human error. Behind it lies a faith in the efficacy of surveillance as a directive force in human affairs. As such, in terms of its understanding of the principal dynamic in the maintenance of clinical standards, it is the ideological opposite of the belief

in the self-sustaining individual professional (or group of professionals) who is motivated by the sense of self-respect that flows from doctors' ownership of professional standards." (Salter 2001: p.874)

Goddard *et. al.* (2000) characterise third way eclecticism as an attempt to strike a balance between oversight, clan, market and network modes of governance. The emphasis on partnership, trust and co-operation signals a move away from competition to network modes of governance. Similarly, clan modes of governance are attenuated by the emphasis on performance indicators under the Performance Assessment Framework and Performance Ratings ('stars'), which signals an attempt to improve information and avoid the excesses of reliance on clan modes of governance.

Taken together, these developments signal a significant step away from market modes of governance associated with competition and contracting, towards new forms of regulatory oversight associated with guidance and monitoring. They constitute a strong system of performance improvement, with an underlying emphasis on the increased accountability of professionals to government, drawing on a comprehensive set of high-level performance indicators and central targets (McLaughlin *et. al.* 2001).

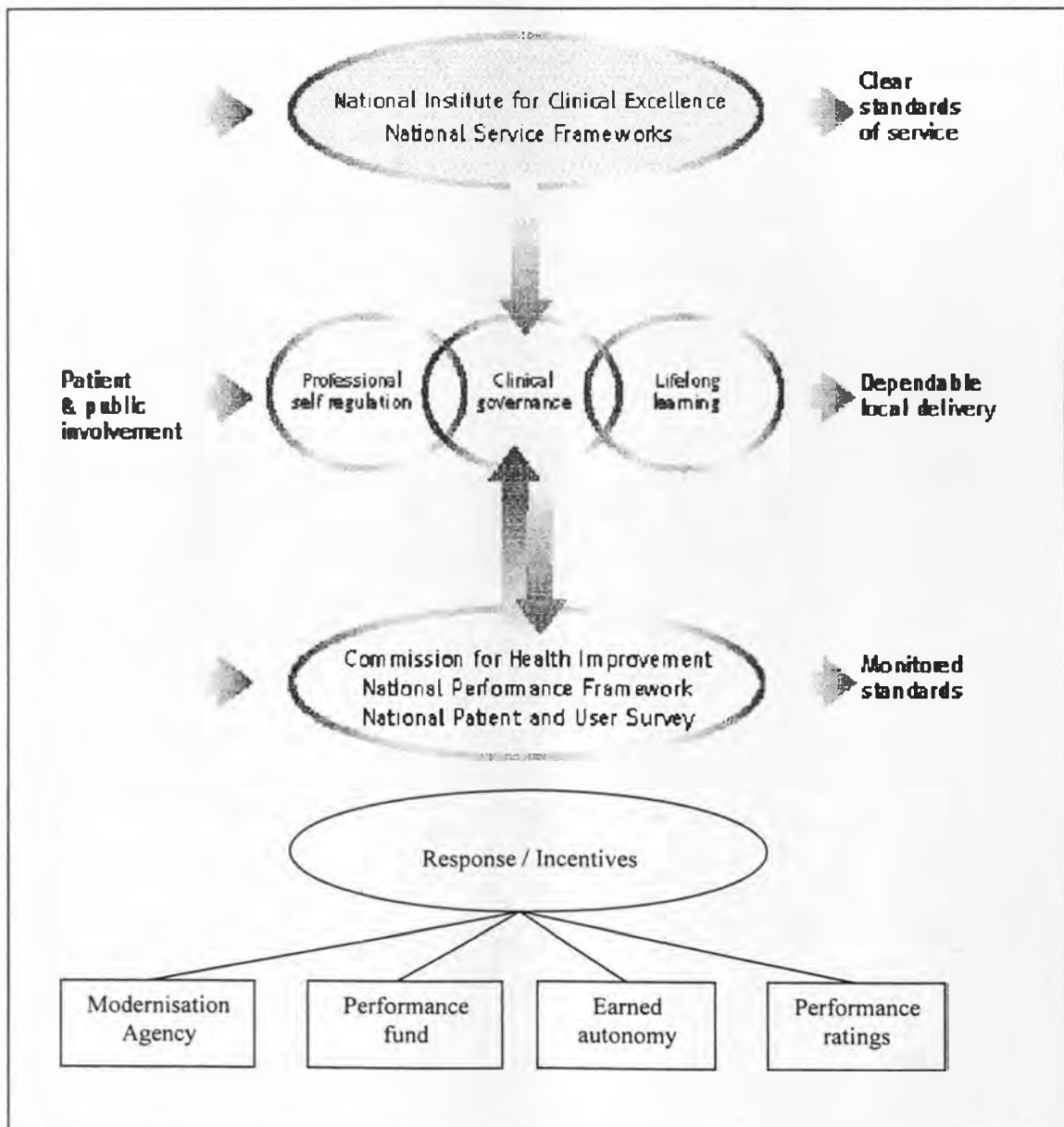
(b) Phase II: The NHS Plan 2000 ('earned autonomy')

Smith (2002) considers developments in the emergent performance management framework contained within the NHS Plan (2000) under three headings: guidance, monitoring and response. These are respectively concerned with transmission of policy objectives; collection of information on the extent to which the guidance has been followed

and targets met; and stimulation of remedial action and / or continuous improvement.

While the guidance mechanisms of NICE and NSFs remained unchanged, there were additions to both monitoring and response systems designed to develop leadership capacity and introduce incentives for improved performance: a new Modernisation Agency; a performance fund; the notion of 'earned autonomy'; and performance rating ('stars') (Figure 1.1.4.3).

Figure 1.1.4.3: The UK health care quality strategy after the NHS Plan, 2000



Launched in April 2001, the Modernisation Agency acts as a focus for developing leadership capacity and local service redesign, focusing primarily on seeking to aid failing organisations. The mechanism for monitoring and assessing organisational performance are performance ratings or 'stars', each trust being awarded between three ('excellent') and zero ('poor') stars depending on their performance against ten key targets, including waiting times, operation cancellations, finance and CHI reports. The performance fund is intended to provide capacity for quality improvement initiatives that enhance progress toward NHS targets. While successful trusts may spend at their discretion, trusts performing poorly under the new performance ratings must spend as directed by the Modernisation Agency. All must report their use of the money, and the intention is to learn from successes and failures for future development purposes.

The third incentive system, earned autonomy, offers local management a degree of autonomy based on organisational performance, including less frequent central monitoring, ability to act as initiative pilot sites, and undertake entrepreneurial activity such take over the management of failing trusts ('franchising'). This incentive system has been extended to allow trusts performing well under the performance rating 'stars' to become foundation trusts with additional freedoms. Notwithstanding the work of the Modernisation Agency in supporting service redesign, Smith (2002) considers the recognition and action upon opportunities for improving performance to be the least developed of the three areas of guidance, monitoring and response.

Alongside national institutions to set, monitor and deliver improved quality, and devolved decision-making subject to earned autonomy Robinson (2002) identifies the selective use of market incentives as a spur to improved efficiency. The 1999 Health Act (DoH 1999) promised greater flexibility to NHS and local authorities in planning partnership arrangements by sanctioning pooled budgets, lead commissioning and integrated provision, all designed to allow greater emphasis on client groups in service design across health and social care boundaries. However:

“Despite claims to the contrary, the emphasis on national standards and accountability set out in ‘delivering the NHS Plan’ suggest that central direction is still an important part of the ministerial mindset. This is likely to place major limitations on the local autonomy elements of the third way-plus model.”
(Robinson 2002: p.24)

As with the earlier experiment of market incentives, the political imperatives for retaining central control may prove to be overwhelming. *Shifting the Balance of Power* (DoH 2001) theoretically gives Primary Care Trusts (PCTs) the devolved power to make decisions over which services are provided and by whom. Similarly, *NHS foundation trusts* (DoH 2002a) outlined increased local freedoms to set local secondary care priorities. Yet, the question still remains as to the extent of their freedom to commission outside of national priorities.

“There is currently an overwhelming emphasis on top-down performance management, with a proliferation of centrally specified targets. But if plans for greater local freedom and earned autonomy materialise, most notably through the creation of foundation trusts, this will be replaced by far greater reliance on bottom-up incentives and drivers.”
(Robinson 2003: p.16)

From the standpoint of early summer 2003, there seems to be an ideological split within the New Labour government concerning the proposed creation of foundation trusts, with supporters claiming increased choice for local people, devolved decision-making and a 'new localism', which will boost performance through increased competition. Critics are wary, concerned about fragmentation and disparities in care quality (Jobanputra and Buchan 2003). Either way, the creation of foundation trusts seems likely to increase the role of market modes of governance within the UK NHS.

1.1.5 The continued rise of the regulatory state?

Walshe (2002) summarises the scope and reach of 'second generation' national regulatory agencies associated with the 1997 reforms and the later NHS Plan (2000) in the English NHS (table 1.1.5.1). These agencies differ from pre-existent bodies such as the Audit Commission or the National Audit Office in their primary focus on regulation, breadth of coverage, and emphasis on clinical quality. As Walshe's emphasis is on regulatory *agencies* he omits three additional regulatory mechanisms: the Performance Assessment Framework (PAF); 'star' ratings; and National Service Frameworks (NSFs), providing information on specific performance indicators, summaries of organisational performance, and expected models of service respectively. Further detail of these measurement mechanisms and their role in assessing progress in clinical governance is explored in chapter 4.

Table 1.1.5.1 'Second generation' NHS regulatory mechanisms: Adapted and extended from Walshe (2002)

Name	Date	Purpose	Methods
National Institute for Clinical Excellence (NICE)	Apr 1999	Provide authoritative guidance on current best practice	Expert review of health technologies, to produce and disseminate guidance
National Service Frameworks	Apr 1999	Provide authoritative advice of the level and type of service provision that patients should expect from specific services, to standardise service provision.	Set national standards of care for specific conditions or services. Publish frameworks and audit assessment tools
Performance Assessment Framework	Apr 1999	To monitor the progress of health care organisations in six performance domains: health improvements, efficiency, fair access, care outcomes, patient experience and effective delivery. Ties service priorities to Treasury Service Agreements.	Collection, assessment and publication of indicator data
Commission for Health Improvement (CHI)	Nov 1999	Assist in addressing unacceptable variations in care	Review all NHS organisations every 4 years; monitor implementation of NICE guidelines and NSFs; investigate systems failures
Modernisation Agency	Apr 2001	To assist in improving service quality and contribute to meeting performance targets	Provision of training and support services to encourage service redesign initiatives, targeted on those areas of service requiring development
National Clinical Assessment Authority	Apr 2001	To support health authorities and trusts faced with concerns over the performance of an individual doctor	Providing advice, taking referrals and undertaking assessments
National Patient Safety Agency	July 2001	Collate and analyses adverse events in NHS; learn lessons and feed them back; produce solutions, set goals and track progress	Operate a mandatory reporting system and provide national guidance/ leadership on safety
Performance Ratings ('Stars')	Sept 2001	Introduce financial incentives to the PAF through development funding and potential to apply for foundation status- 'earned autonomy'	Summary measure of achievement against key targets. Assessment affect access to service development monies.
Commission for Health care Audit and Inspection (CHAI)	Apr 2004	To help improve the quality of patient care by assisting to address variations and ensure high standards of patient care	Reviewing NHS organisations on a rolling visit programme; overseeing Performance Ratings and indicators under the PAF; Monitoring NSFs and NICE guideline implementation.

The need for a balance between trusting professionals to develop high quality services and checking up on them to ensure that they do so is widely accepted (Davies et. al. 2001). The evolution of the Commission for Health Improvement (CHI) to the Commission for Health Care Audit and Inspection (CHAI) could be seen as an attempt to rationalise the agencies involved in promoting clinical governance. Under these new arrangements, the modernisation agency has the lead on building trust and capacity for improvement, CHAI taking the lead on checking via performance ratings and visits. However, the tensions and resultant pressures on those delivering services remain.

Opinions on the desirability of the growth of regulation within the UK NHS differ. Davies Mannion & Marshall (2001) are broadly supportive of the balancing of mutuality, oversight and competition in the NHS Plan reforms, and Walshe (2002) summarises the managerialist perspective well:

"The rise of regulation in the NHS seems, at first sight, to represent a long-term strengthening of central government's control of managerial practice... Their creation could allow a kind of centralised micromanagement, in which there is less and less scope for local variation to develop. However, if the politicians can be persuaded to let go, the new regulators of the NHS could provide a genuinely new approach to improving performance and management, [providing] a more indirect and distanced relationship, in which managed regulation by intermediate and quasi-independent organisations would play a large part. In that environment, the new regulatory agencies could adopt a more responsive approach to regulation. "
(Walshe 2002: p. 969)

Yet, such optimism is not universally shared and the conditional statements in the fourth and fifth line may well prove to be telling. More fundamentally, critics of New Labour's

conception of policy implementation point to a dramatic expansion of treasury control over policy outputs, in terms of performance criteria for service delivery enshrined in the PAF (Lee & Woodward 2002). They characterise the third way approach to policy implementation as increased central policy prescription through treasury resource allocation tied to service agreements. In this context, earned autonomy as introduced under the NHS Plan (DoH 2000) means the delegation of responsibility for administering national priorities- it is administrative rather than political devolution. Together, treasury service agreements and administrative responsibility for delivery have effectively 'nationalised' policy implementation:

"By attempting to both steer and row through greater central prescription by the core executive, the Third Way of implementation has denied inter-organisational networks at sub-national levels of governance the necessary autonomy to innovate or show foresight." (Lee & Woodward 2002: p. 53)

Thus those emphasising network modes of governance within the clinical governance project see clinical governance as an opportunity (Walsh & Small 2001). In contrast, those concentrating their attention on oversight mechanisms such as the PAF characterise clinical governance as creeping central control (Harrison & Lim 2000; Lee & Woodward 2000; Miles *et. al.* 2000). While Ham (1999b) is correct to chide those who characterise the New Labour reforms as simply increased central control, critics do make a strong case. Each reveals a partial truth: clinical governance is not itself a mode of governance, but draws on hierarchical, network and clan modes of governance. The concern is that while there is some scope for administrative autonomy in service delivery, there is arguably an

overemphasis on external summative approaches to the extent that the internal, formative, improvement agenda is overshadowed.

1.2 The importance of measuring progress in clinical governance

The balance struck between checking and trusting in modes of governance reveals assumptions about the motivations of managers and clinicians and has important implications for staff motivation. While the third way combines planning, competition and new forms of regulation in a complex blend of levers the tensions are arguably strongest between techniques compelling performance improvement through externally applied measurement (oversight), and approaches that trust and support intrinsic professional motivation to deliver high quality services (mutuality). Ham (1999a) acknowledges this tension, cautioning against the assumption that NHS staff are not fundamentally well motivated:

“No amount of guidance from the NHS Executive or hectoring by politicians can substitute for a drive to improve performance that comes from within and is acknowledged and valued by those steering the process of change.”
(Ham 1999a: p. 1492)

Consequently robust oversight mechanisms such as those outlined within the 1997 White Paper and NHS Plan (DoH 2000) may be counter-balanced by developing a quality-focused and learning –oriented culture, in order to re-employ the strong internal motivations of staff

and reduce agent waywardness. It is for this reason that the clinical governance literature so strongly emphasises an open organisational culture as a medium-term objective.

Trusts face a difficult balancing act as they seek to align central and local agendas for quality improvement dynamically, while meeting their performance measurement requirements (Degeling, Hunter & Dowswell 2001; Lawton, McKevitt & Millar 2000). Validated measures of staff perceptions of progress could provide trusts with valuable information on staff perceptions on clinical governance climate, which they may use to identify priorities for organisational development activity and assess their efficacy over time. Currently, the danger is that given the policy context, the pursuit of the long-term agenda may leave organisations open to accusations of not having robust enough performance management. Use of such measures may provide trusts with valuable evidence of a robust, long-term focus on quality improvement, with which to reassure external agencies. Most importantly, they would also highlight areas requiring attention to secure the required cultural attributes of openness, continuous improvement and proactive risk management.

1.3 The contribution of the thesis

The thesis reviews published evaluations of progress in clinical governance, develops two validated measures of progress in order to fill the gap in available instrumentation, and via their application examines the effect of tensions within clinical governance on policy

implementation. Chapter two reviews the empirical literature on clinical governance implementation to date in order to identify perceived successes and areas of difficulty, as well as highlight gaps in available approaches to measurement. Chapter three explores the literature on use of performance indicators for internal quality improvement and external quality assurance, identifying good practice and potential pitfalls to avoid in indicator design. As a precursor to the design, validation and use of two new measures of progress in clinical governance, chapter four identifies a range of approaches available to measure progress in clinical governance, making a case for the development of two additional measures.

The development and validation of the first of these two measures, the Clinical Governance Climate Questionnaire (CGCQ), is outlined in chapter five. The CGCQ is designed to assess the climate of clinical governance within organisations at the level of individual staff experience. Following development, the results of a study of the CGCQ involving over 400 respondents across four NHS trusts are reported and discussed in chapter six. The second measure, the Organisational Progress in Clinical Governance (OPCG) schedule, is designed to assess organisational progress across the range of competencies required for clinical governance, administered at board and directorate level. Chapter seven details the development, validity and reliability testing of the OPCG. Chapter eight reports on a national study of progress in clinical governance using the OPCG based on 1170 respondents across 100 NHS trusts, providing an overview of achievements to date across all measurement domains of clinical governance.

Thus, the thesis makes both conceptual and empirical contributions to the object of study. Conceptually, it explores the links between clinical governance and broader governance trends, identifying internal tensions that have influenced implementation (chapters one and two). Current knowledge concerning the application of performance indicators to public sector bureaucracies is synthesised in chapter three. Chapter four identifies the paucity of internal, formative measurement indices of clinical governance progress and makes the case for measures of organisational climate and competency, exploring conceptual difficulties involved in the measurement of organisational climate. Chapter five inductively generates a conceptual map of clinical governance climate, drawing on prior literature, qualitative exploration of potential measurement domains and an exploratory factor analysis of the derived indicators. The conceptual difficulties involved in levels of aggregation in measures of climate are explored in chapter six. Chapter seven provides a conceptual exploration of competency frameworks in the context of organisational progress in clinical governance.

Empirically, the thesis provides two validated measures of clinical governance progress for use at different levels of aggregation: with individual staff (CGCQ); and at directorate manager and board level (OPCG). In addition to the development, validation and reliability (internal consistency) of both measurement tools (chapter five and chapter seven), results are reported of their implementation (chapters six and eight). Chapter eight explores the effects of the internal tensions within clinical governance on implementation, drawing on empirical results of a national sample of trusts using the OPCG.

Conclusions

Internally, NHS organisations are obliged to develop systems of clinical governance to ensure that they are delivering high quality clinical care, consistent with standards established nationally through the National Institute of Clinical Excellence (NICE) and National Service Frameworks (NSFs). Internal mechanisms include clear lines of accountability to the board and a committee structure, and statutory responsibilities for CEOs. NHS organisations are externally supported by the Commission for Health Improvement (CHI), who intend to visit every healthcare organisation on a rolling four-year visit programme to see that clinical governance arrangements are robust.

Organisations scoring well on indicators contained in the Performance Assessment Framework (PAF) will be subject to less scrutiny and receive money for service development, the notion of 'earned autonomy'.

The New Labour NHS reforms address accountability by increasing levels of internal and external scrutiny, while simultaneously reducing and reformulating the role of professionalism within the context of safeguarding clinical quality. The new central agencies, concerned with setting and monitoring clear service standards, provide the context. The role of local clinical governance arrangements within the system is to ensure 'dependable local delivery' of these standards, and provide information to support assurance internally and externally. Self-regulation within this system needs to be 'rigorous' i.e. tied firmly to the local delivery of national targets, rather than an ethical code and cornerstone

of professional autonomy. From this perspective, clinical governance represents a new balance between oversight ('checking') and mutuality ('trusting') modes of governance in service accountability.

Yet the twin objectives of assurance and improvement, related to oversight and mutuality modes of governance respectively, result in significant tensions. Clinical governance requires the delivery of a central agenda on one hand, while nurturing devolved continuous quality improvement on the other. These two aims require different sorts of information, for use in different ways, by different groups of people, and consequently they are not easy to run in parallel. Within devolved provider organisations the temptation to game performance indicators is very high, given the potential rewards of reduced scrutiny and funds for improvements given to those scoring well. Further, there is empirical evidence that providers may emphasise performance management at the expense of the quality improvement agenda (Freeman *et. al.* 2001), which is understandable in an environment in which the CEO's job and development monies are dependent upon attainment of performance targets set by central government.

The tension between the needs of assurance and improvement is reflected in the difference between the short- and medium-term targets for clinical governance outlined in the discussion document (DoH 1998). While the medium term agenda emphasises the need for cultural transformation including greater openness and avoidance of unnecessary blame, the short-term targets relate to committee structures. There is a danger that as the structural work is completed, organisations feel that they are now 'doing' clinical governance, leaving

the long-term quality improvement agenda underdeveloped. Given the importance of the conceptual disputes concerning clinical governance, chapter two reviews evaluations of its implementation to assess the extent of progress in quality assurance and quality improvement.

CHAPTER 2

THE EARLY DEVELOPMENT OF CLINICAL GOVERNANCE: A REVIEW OF EMPIRICAL EVALUATIONS

"Quality improvement cannot be imposed by decree."
(Campbell *et. al.* 2001a: p. 1582)

Introduction

This chapter reviews the empirical literature on the early development of clinical governance, identifying broad trends and issues in the translation of policy into practice. Achievements, areas of concern and potential future difficulties are explored across primary and secondary health care sectors. While there are sector-specific difficulties, a careful reading of the literature reveals a number of implementation difficulties common to both. These common difficulties are occasioned by internal tensions within clinical governance, principally the combination of quality improvement and quality assurance. Common implementation difficulties notwithstanding, the divergent contexts, histories and traditions of primary and secondary care provision have subtly influenced the way that these difficulties have been addressed. Within primary care a greater emphasis on educative and cultural change dimensions of clinical governance may be discerned, while the secondary sector exhibits a contrasting emphasis on structures and performance management systems. Commentators highlight different aspects of clinical governance, some concentrating on achievement against statutory targets, others on structural integration or implementation barriers. Consequently, there is no single comprehensive overview of progress in all of the domains of clinical governance. Further, the effects of change are almost exclusively

explored through the views of those charged with implementing clinical governance, and there are as yet no validated instruments for assessing the 'fit' between clinician attitudes and those one might expect to find were clinical governance successfully integrated into everyday practice.

2.1 Methods

In order to identify issues in the early implementation of clinical governance, a literature review of empirical studies of its development in primary, secondary and tertiary health care was undertaken. Studies reporting empirical results were included if they related to multi-site evaluations of progress in clinical governance, or identified potential future implementation difficulties.

Data Sources, Selection and Data Extraction

The study comprised an extensive review of the electronic databases Medline, HMIC, ASSIA and BIDS selected for their depth of coverage of health care management topics. Searches were limited to articles or reports published between January 1997 and December 2002, using a simple truncated search for 'clinical governance' in combination with 'implement\$' within the document. All identified abstracts were reviewed, and full articles retrieved where it was clear that they contained empirical evidence on clinical governance implementation. Where this was not clear from the abstract, full articles were retrieved before making a decision on inclusion or exclusion. Given the tight focus, papers detailing

the development of clinical governance structures within a specific unit or organisation were excluded. Studies included cross-sectional and longitudinal surveys with either regional or national focus; semi-structured interviews with board level trust staff; and a content analysis of trust clinical governance reports from year one. Individual study methodologies were not evaluated due to their largely descriptive nature.

Two major programmes of research with multiple publications were identified, one in primary care and the other predominantly in the secondary and tertiary sector. The national tracker survey team at the National Primary Care Research and Development Centre (NPCRDC) undertook the primary care programme. Their series of three cross-sectional surveys in December 1999, 2000 and 2001 provide a longitudinal overview of the development of clinical governance in 72 of the 481 English Primary Care Groups/ Trusts (PCG/Ts). An overview of clinical governance development in secondary and tertiary care is provided by Walshe and colleagues from the University of Birmingham's Health Services Management Centre in a series of census surveys and semi-structured interview studies, primarily based on developments within the former West Midlands and South West Regions.

Data Synthesis

In total, 24 articles were selected. Given the review's focus on the broad themes of good progress, areas for concern, and potential future difficulties in clinical governance implementation, data extraction techniques were designed to facilitate the exploration of each theme and the inductive generation of sub-themes. Consequently, each article was

carefully read to identify key findings under each theme, which were noted and recorded in a matrix (appendix 2). The matrix may be read either across (rows) to identify information relating to each of the themes within a specific article; or down (columns) to identify information relating to each of the articles on a specific theme. Sub-themes were generated inductively for each theme by reading the appropriate matrix column in sequence and summarising the major elements, and are reported below in narrative form. Given the different traditions of primary care and secondary / tertiary services results are reported separately.

2.2 Results

2.2.1 Primary care

"The approach of the UK government [to clinical governance in primary care] is heavily top down, with detailed processes and guidelines being determined at national level. For these to be successful, significant cultural change is necessary for clinicians on the ground to engage with the quality improvement agenda. It is as yet unclear whether the right mix of national guidance and local freedom is being provided."
(Roland *et. al.* 2001)

Established in England in April 1999 as the organisational framework for delivery of primary care services, Primary Care Groups (PCGs) are also responsible for implementing clinical governance. Initially operating as subcommittees of the then Health Authorities, PCGs are working towards Primary Care Trust (PCT) status, in which they hold their own budget for the health care of their population, by April 2004 at the latest. This marks a major shift in the provision of primary care within England. Previously, general

practitioners tended to work within their own practices rather than as part of a larger organisation, and consequently primary and community services tended to be fragmented. The transition to a more integrated framework implied by clinical governance poses a major challenge, as it requires the development of a corporate culture for shared quality improvement activity, in the context of large-scale organisational change.

Good progress

Notwithstanding the extent of the challenge, empirical assessments of early implementation found evidence of good progress on a number of fronts.

Baseline and infrastructure

PCG/Ts were quick to undertake an assessment of quality assurance and improvement capacity and develop local implementation strategies. An early national survey of PCG/T clinical governance leads (Preston & Baker 2000) identified that 96% of respondents had developed or were developing a clinical governance strategy as a result of a baseline assessment of infrastructure for clinical governance. While this seems impressive, the thoroughness with which these baseline assessments were undertaken is less clear, as is the extent of the restructuring occasioned by the reviews. An exploratory factor analysis of emergent clinical governance strategies identified seven factors: risk management; IT strategy; clinical audit; external benchmarking; quality and roles; access to information; and training and development (Spurgeon & De Luc 2002)

Shared learning and partnerships

At the outset, a national survey of clinical governance leads identified a number of pressing educational needs (Preston and Baker 2000), including change management (157, 56.9%); leadership (133, 40.5%); and facilitation skills (130, 39.5%). While daunting, there is evidence that PCG/Ts made considerable progress in use of educative approaches to quality improvement. By December 2000, 93% of respondents to the NPCRDC national tracker survey were implementing practice and personal learning plans for GPs, compared to only 4% the previous year (Campbell *et al.* 2001a). At this time, the most common inter-organisational activity was protected learning time (81%), typically provided via monthly half-day learning events, for either GPs or the whole primary care team. Reported attendance rates at events were high, often over 95% (Roland *et al.* 2001). Perhaps most impressively, activity was not limited to initiatives within organisations; there is evidence of a range of partnership working between different local health care agencies, including hospital trusts (60%); other PCGs/Ts (57%); and community health services (38%) (Campbell *et al.* 2001a). Such activities represent a major change for UK general practice; GPs were far less likely to be working in isolation, and typically had access to facilitated learning and opportunities to engage in quality improvement activity with a wide group of peers (Roland *et al.* 2001). This represents an opportunity for the broader primary care team to be involved, providing a multi-disciplinary forum for quality improvement.

Sharing information

Historically, information on care quality was often not shared with others within individual practices, and examples of shared information between practices were exceedingly rare. By

December 2000, practices were increasingly making information on care quality available, aggregated at both PCG/T and individual practice level, to board members (62%), general practices (50%) and even the public (41%) in anonymised form (Campbell *et. al.* 2001a). The trend for wider sharing of information continued; by December 2001, 81% of practices were sharing information on quality. This marked a significant change in the desirability of information sharing with both professional peers and the public.

Cultural change

While provided largely by interviews or surveys with those responsible for clinical governance implementation and thus to be viewed with caution, there is evidence of substantial activity within primary care likely to result in changes to the cultural climate of quality improvement. Activities were undertaken with an emphasis on building sustainable infrastructures in audit / risk management and guideline activity, and implement 'bottom-up' quality improvement initiatives such as the sharing of anonymised data for benchmarking. Essentially, the focus has been one of developing infrastructures and supportive cultures for quality improvement, rather than upon specific target areas or performance management frameworks- facilitative non-policing approaches (Campbell *et. al.* 2002). While patchy in coverage, considerable amounts of quality improvement work have previously been undertaken in primary care, especially in teaching practices (Taylor 2000). Consequently, PCG/Ts have tended to build on pre-existent quality improvement activity, seeking to develop greater corporacy across all practices and building trust by sharing information (Roland *et. al.* 2001; Campbell *et. al.* 2002).

Use of financial and contractual incentives for quality improvement

While the thrust of developments in primary care was largely educational, increasing use was made of financial and contractual incentives to support the desired changes.

Consequently, the proportion of PCG's/Ts using financial and contractual incentives to focus attention on quality improvement targets, rose steadily from 29% in December 1999 to 50% by December 2000 (Campbell 2001b), and 65% by December 2001 (Wilkin *et. al.* 2002). The increase in use of financial incentives was also found by Spurgeon and de Luc (2002), 88% of their respondents citing financial incentives as a lever.

Areas causing concern

Structural, cultural and resource barriers to implementation may be discerned (Campbell *et. al.* 2002).

Structural barriers

These include a lack of contractual levers to influence general practices, other than Personal Medical Services (PMS) practices in which GPs are directly employed. Independent contractual status means that levers such as financial incentives offer greater potential for influence. The comparative lack of performance management mechanisms available in primary care in relation to secondary / tertiary services may partially explain the preference for supportive and developmental approaches within primary care, and the year-on-year rise in the use of financial incentives as a lever; currently, mechanisms for quality assurance are weak. This may change should the new GMS contract be agreed, given the increased

provision of performance indicators under the contract. Only 20% of PCG/Ts were using disciplinary procedures to deal with poor performance (Wilkin *et. al.* 2002). Further structural problems include the highly variable stages of development of different practices in terms of commitment and infrastructure, exacerbated by a fast-changing organisational environment.

Cultural barriers

While there is evidence of some development of supportive cultures for quality improvement, cultural barriers such as a perceived blame culture associated with quality assurance approaches are difficult to address and will require long-term ongoing work to reassure and involve clinicians. Major identified barriers included a history of GP independence making joint work difficult, the extent of political imperative around specific initiatives, and a perceived lack of skills in change management (Marshall *et. al.* 2002).

Resource restrictions

In December 1999, 85% of PCG/Ts had no support from the Health Authority (Campbell *et al.* 2001b); by December 2002 42% of PCG/Ts still had no implementation budget and hence no way of paying for dedicated support staff (Wilkin *et. al.* 2002). Such financial restrictions pose a serious threat to the continued development of clinical governance in this sector.

Potential future difficulties

Speed of change

Campbell *et. al.* (2001a) point to the lack of experience of involving large groups of primary care staff in ongoing quality improvement activity, identifying difficulties encountered in the introduction of more limited schemes, such as lack of authority to implement change, unclear roles and responsibilities, and lack of information on care quality. They caution against driving the pace of changes too fast and disengaging the very people whose attitudes and behaviours the policy is designed to influence.

Lack of hierarchical structures for enforcing bureaucratic modes of governance in PCTs

The organisational context of quality improvement in UK primary care poses a particular set of difficulties. PCTs are run by Professional Executive Committees (PECs), responsible for policy and its implementation, overseen by a trust board with a majority of lay members. Unlike their hospital counterparts, these primary care structures do not have direct managerial lines of accountability, thus the tension between the desire of clinicians to be self-regulating and the aim of the government to manage the system is likely to be particularly acute (Roland *et. al.* 2001). Harrison & Lim (2000) conceptualise the difficulty as one of a mismatch between bureaucratic modes of governance (control of professionals via a web of rules) within a largely collegial context, unfitted to implementing bureaucratic rules. This may partially explain the extent to which PCG/Ts seemed committed to bottom-up development and securing clinician support- it provides a pragmatic means to increase

the importance of clinical quality in a collegial environment with weak enforcement mechanisms.

However, there are difficulties with the above analysis. Firstly, it may present an inaccurate picture of clinical management within the acute sector. Medical directors may well be formally accountable for their medical colleagues, but are required to discharge their responsibilities with professional colleagues, requiring persuasion and influence rather than diktat. The collegiate nature of clinician influence sits uncomfortably with accountability functions (Freeman *et. al.* 1999; Latham *et. al.* 1999). Further, Walsh & Small (2001) are critical of characterisations of clinical governance as bureaucratic imposition. For them, meaningful clinical governance requires the building of self-organising networks from the bottom-up, in which space is afforded for continuing professional judgements; clinical governance may be an opportunity rather than an imposition. Thus the disputed nature of clinical governance - as external assurance or internal quality improvement - again influences the debate.

2.2.2 Secondary care

“In only a few trusts did the [trust annual clinical governance] report make explicit the realisation that the issue was how to provide quality care by engaging clinical staff...”
(Franks 2001)

In the first year of implementation (April 1999 – April 2000) acute, tertiary and ambulance trusts were charged with making progress in four areas: assigning leadership, accountability

and working relationships; undertaking a baseline assessment of capacity and capability; formulating an agreed plan; and clarifying reporting arrangements for clinical governance (NHSE 1999). Given such a strong lead, early activity was largely centred on these elements.

Good progress

Leadership and accountability

Latham *et. al.* (2000) reported that almost all (70, 95%) trusts in the West Midlands and South West regions had identified an executive director of the board with specific responsibility for clinical governance by the end of 1999. Of these, almost all had designated a senior clinician as clinical governance lead, suggesting that leadership for clinical governance was widely regarded as a clinical responsibility (Freeman *et. al.* 1999). For their part, clinical governance leads were both positive and committed to their roles (Walshe *et. al.* 2000).

Baseline assessment and strategic development

The guidance on clinical governance implementation clearly stated that strategic development should proceed from a critical baseline assessment of current structures and process for quality assurance and improvement, in order to clarify the precise needs of the organisation. With several months still to go before the April 2000 deadline, 49 (66%) of trusts in the West Midlands and South West regions had undertaken a baseline assessment of current systems, 44 (59%) had produced a strategy, and 27 (36%) had developed an action plan (Latham *et. al.* 2000). Key areas identified in action planning included

strengthening existing systems for quality improvement, professional development and appraisal.

Reporting arrangements: committee structures

NHS trusts rapidly established sub-board clinical governance committees, indeed by October 1999 all trusts in the West Midlands and South West Regions had done so (Freeman *et. al.* 1999; Latham *et. al.* 1999). However, the existence of a structure is only a first step, as to successfully discharge their function they require clarity of purpose, and an appropriate membership. While most (53, 72%) committees had written terms of reference, some ambiguity over the proper purpose and function(s) of such committees is evident from differences in size, membership composition, reporting relationships and meeting frequency across trusts (Latham *et. al.* 2000). Memberships ranged from three to 23, the larger committees including non-executive directors and representatives from primary care, health authority and community health council (CHC). Such variation is most likely due to differences in committee remit, smaller committees suited to a monitoring role, and larger committees suited to an operational role to secure wide support (Latham *et. al.* 2000). While infrastructure at the corporate level is well documented, there is little information available on infrastructure at the level of directorates or clinical teams. Consequently, it is difficult to assess the degree to which trusts were able to co-ordinate clinical governance activity within clinical areas and monitor their progress (Franks 2001).

Use of education and training initiatives in implementation

In terms of organisational interventions to implement clinical governance, educational initiatives predominated. These largely consisted of education and training exercises aimed at awareness raising, but also included evidence-based practice development and clinical training (Wallace *et. al.* 2001a; Walshe *et. al.* 2000; Wallace *et. al.* 2001b).

Areas causing concern

While resource constraints of time and money will always pose problems, further potential difficulties included the availability of timely and accurate clinical data, poor existing infrastructure for quality improvement and a relatively superficial restructuring of infrastructure for clinical governance.

Poor clinical record keeping

Concern was expressed over the quality of clinical record keeping (Conduit *et. al.* 1999) both in terms of the accuracy of the content and their effectiveness as a communication tool between team members. Without such information, it is difficult to assess current clinical effectiveness, or to evaluate the outcomes of quality improvement activity on clinical quality.

Variability of coverage and effectiveness of existing quality improvement and assurance systems

Formal guidance on the development of clinical governance emphasised the importance of a comprehensive programme of quality improvement activities (DoH 1998; NHSE 1999), including clinical audit, risk / complaints / adverse event management, professional development and knowledge management. Notwithstanding over ten years of mandatory quality improvement activity within secondary care in the UK NHS, organisational arrangements for existing quality improvement systems were highly variable in coverage and perceived effectiveness both within and between trusts (Walshe *et. al.* 2000).

A postal census of all West Midland and South West region clinical governance leads (Freeman *et. al.* 2001a) found that while a majority of trusts had established groups to oversee clinical audit, risk management, complaints and adverse incidents, rather fewer had such structures for CPD and knowledge management. When asked to indicate how well existing systems for quality improvement covered directorates or service areas, 51 (73%) respondents believed their systems for complaints and adverse incidents and clinical audit covered most (over 75%) service areas. In contrast, only 6 (10%) trusts' systems for knowledge management achieved this level of coverage (Freeman *et. al.* 2001a). The perceived effectiveness of systems was also highly variable; while over 70% of respondents believed that systems for complaints and adverse events were effective, less than 30% considered arrangements for clinical audit, risk management or professional development to be effective.

Superficial nature of much restructuring



While there is little doubt that there was considerable room for improvement in both the coverage and effectiveness of quality improvement systems, there is some evidence that the mandatory imperative of baseline assessment and action planning tended to produce less thorough reviews than one might hope. Based on a content analysis of trust annual reports, Franks (2001) outlines a typology of three possible approaches to organisational restructuring occasioned by clinical governance. The least sophisticated approach is that of tactical *structuralists*, who simply retained pre-existing organisational elements, realigning them within a hierarchical structure to seemingly meet the new requirements. More thorough than tactical structuralists, *functionalists* considered the functional elements required, creating new one where they are not already in existence, and aligning the elements hierarchically in a new structure. Most sophisticated of all, *strategists* considered the issue holistically, addressing their fundamental approach to quality within an explicit model, such as European Foundation for Quality Management Excellence Model (EFQM EM).

The strategic approach to clinical governance implementation was rare (Franks 2001). The superficial nature of much restructuring and strategic planning is supported by Latham *et al.* (2000), who report that of 44 trusts who had produced a clinical governance strategy, 15 (34%) had done so without undertaking a prior baseline assessment. It is within the context of a strategic approach that organisations are most likely to explicitly address the issue of behavioural change as part of organisational development. Consequently, the rarity in use of models such as the EFQM EM may in part explain the rarity of strategic planning for organisational development in clinical governance implementation.

Clinical leadership

While there is evidence that almost all trusts had appointed an executive director of the board with specific responsibility for clinical governance, the amount of time spent by such individuals varied greatly, between none and 50 hours per week (Latham *et. al.* 1999; Latham *et. al.* 2000); Freeman *et. al.* 1999). Most (55, 74%) had no time specifically allocated for the role (Latham *et. al.* 2000), and it is difficult to imagine that responsibilities for clinical governance could be discharged effectively without protected time.

User involvement

The extent of user involvement in quality improvement was considered poor, and meaningful involvement somewhat lacking (Dewar 1999; Franks 2001)

Frequency of reporting for corporate accountability

The test of senior commitment to clinical governance is to be found less in rhetoric or structural development than in consistency of action (Latham *et. al.* 2000), partly discernible by the frequency of board-level reporting on clinical governance issues and consequent actions. There is some evidence of scope for increased board-level reporting of clinical governance issues (Latham *et. al.* 2000).

Potential future difficulties

Major potential difficulties included a lack of planning for organisational development (Wallace *et. al.* 2001a; Franks 2001), a deficit of transformational leadership styles

(Wallace *et. al.* 2001b), and assurance focused measurement systems for monitoring implementation (Freeman *et. al.* 2001b).

1. Lack of planning for Organisational Development (OD)

Drawing on a range of disciplines within the social sciences including systems theory and theories of behavioural change, Organisational Development (OD) refers to approaches used to influence behaviour and culture within organisations. In an assessment of the application of OD techniques with clinical governance implementation, Wallace *et. al.* (2001a; 2000b) explored the extent to which executives and non-executives had a clear understanding of the likely impact of clinical governance on organisational cultures and their role in achieving their local objectives, from an OD perspective. They found little evidence of change management plans identifying the means of achieving attitudinal change in specific sub-groups such as managers, doctors, or other clinicians (Wallace *et. al.* 2001a). While the overwhelming majority of respondents identified education and training events aimed at awareness raising, this did not amount to a strategy for change management. Rather, trusts were using what amounted to multiple optional means, anticipating organisation-wide change on the basis of unspecified diffusion processes (Wallace *et. al.* 2001b). Further evidence of a lack of strategic planning was provided by a strong relationship between the planned future use of interventions and past use, rather than perceived effectiveness i.e. the choice of intervention seemed to be influenced more by historical precedent than what was perceived to work.

2. Lack of transformational leadership styles

Despite a widespread view that a more open culture around reporting was desirable, there is little evidence that executives were explicitly taking responsibility for leading change (Wallace *et. al.* 2001a). Leaders were failing to take a systematic approach to the design and implementation of interventions that could lead to positive culture change. They did not have a clear vision of their intended outcomes, were not assessing and monitoring cultural changes, nor did they explicitly plan to influence the attitudes and behaviours of staff (Walshe *et. al.* 2001).

3. *Assurance focused measurement systems for monitoring implementation*

The tension between quality improvement and quality assurance within clinical governance was reflected in the intentions of trust directors towards monitoring implementation (Freeman *et. al.* 2001b). Differences in measurement intentions revealed differences in emphasis on mutuality, competition and oversight as modes of governance, as explored in chapter 1. Those directors who emphasised the role of quality improvement within clinical governance proposed measures of culture, OD and staff involvement; those who emphasised the assurance function proposed measures of systems, processes, outcomes and lines of accountability. While the variety and scope of potential identifiers was heartening, and included measures of organisational culture, organisational development, and staff engagement in change, only 48 (36%) interviewees intended to incorporate elements of cultural change and organisational development into measurement systems. The majority (86, 64%) described their intentions wholly in terms of assurance.

To rely exclusively on such approaches may prove mistaken. It is likely that the aims of a performance management system will be compromised unless counterbalanced by a quality-focused climate, and the involvement of staff in the design of indicators of cultural change may help to offset the tendency of measurement, monitoring and control to develop unintended consequences. These issues are explored in further detail in chapter 3.

2.2.3 Shared potential future difficulties

While each sector faces specific difficulties occasioned by their context, they share a number of implementation difficulties as a consequence of the conceptual tensions within clinical governance itself. Principally these comprise the importance of both developing a non-blame culture and supporting quality improvement, while at the same time dealing with poor performance and providing accountability for clinical quality via quality assurance.

1. Engagement with QI / dealing with poor performance

PCG/Ts have taken a largely educational approach to deal with poor clinical performance, strategies including informal discussions, providing training and allocating additional resources to poorly performing practices. The performance management role of Health Authorities made it possible for PCGs to concentrate on the formative agenda, supporting clinicians to develop services. The move to PCT status and demise of the Health Authority tier in April 2002 poses difficulties for the future balance between formative and summative approaches within primary care. While an emphasis on educative approaches

remained strong, PCG/Ts increasingly used formal disciplinary procedures, rising from 9% in 2000/2001 to 20% in 2001/ 2002 (Wilkin *et. al.* 2002).

"Many of those responsible for clinical governance told us that in order to engage health practitioners in quality improvement, it is essential for them to be seen as helpful to and supportive of practices. However, this may become more difficult as groups become primary care trusts, a move that will ensure that they take on more responsibility for the quality of care provided by clinicians in their area." (Campbell *et. al.* 2001a: p. 1582)

There is as yet little empirical evidence on the impact of this shifting balance on the ability of PCG/Ts to engage clinicians in the quality improvement agenda.

2. Quality improvement / quality assurance

Across primary, secondary and tertiary care, early implementation of clinical governance incorporated a wide variety of approaches drawing on systems for assuring existing care quality and improving future care quality. While this is unsurprising given the explicit intent of clinical governance to combine these two approaches (NHS Plan 2000), the tension between them has important implications for implementation. The former refer to accountability mechanisms by which the government is able to ensure that clinical care reaches minimum standards, and that poor performers are identified and dealt with. In contrast, the latter aims to incorporate activities to improve quality into everyday activity- to prevent problems and design them out of processes rather than inspect practice to identify shortfalls after the event.

Organisations may face pressure from above and below, and in seeking both assurance and improvement may achieve neither (Campbell *et. al.* 2002). On the one hand, despite an

emphasis on development and support, the quality assurance element may foster suspicion and disengagement on the part of some clinicians; on the other, supportive and educational strategies may not satisfy government desires for increased accountability and conformance with standards. The tension has been characterised as a serious flaw in the implementation of clinical governance (Roland *et. al.* 2001), given the potential for hostility from clinicians to perceived surveillance of their clinical activity. Consequentially, Roland *et. al.*(2001) counsel for a counter-intuitive 'light touch' assurance regime, concentrating activity on continuous improvement. While at first glance counter-intuitive, such an approach would encourage the involvement of clinicians. However, the strategy may prove difficult in the face of strong performance management against a clearly defined central agenda.

3. Non-blame culture / tackling poor performance

The PCG/T experience has largely been to emphasise the educative, internal and formative elements of clinical governance - supporting practices and encouraging multi-professional service development- rather than the external, summative elements, in an attempt to minimise blame culture and ensure professional engagement (Campbell *et. al.* 2002).

Within secondary care, the emphasis has been on development of appraisal systems for clinical, and especially consultant medical, staff. This has required appraisers to delicately balance the needs of the organisation to assure care quality with the desire of professionals to continually update their practice (Walshe *et. al.* 2000). Reconciling the need for a no-blame culture while simultaneously tackling poor performance remains a key issue for both primary and secondary health care organisations.

4. *Lack of evidence that clinical governance has improved quality of care or 'how things are done'*

Notwithstanding evidence of much change in structures and processes for clinical governance, there is little evidence that such activity has improved clinical outcomes. Several commentators remark that this may be due simply to the lack of elapsed time since implementation began (Roland *et. al.* 2001; Walshe *et. al.* 2000), and that it would be premature to expect to see quantifiable improvements so quickly. During the early period there were no empirical assessments of the impact of the policy on clinical outcomes, due partly to methodological complexities pertaining to attribution of effect in the context of mandatory policy.

Conclusions

Empirical studies reveal a great deal of positive development throughout the early years of clinical governance implementation in primary care. The first year (1999/2000) saw a concentration of activity on baseline assessment and developing an infrastructure. There is evidence of considerable progress in securing the commitment and involvement of primary care teams in 2000/2001, facilitated largely by a supportive and educative approach to implementation. This was further developed in year three (2001-2002), as PCG/Ts continued to develop supportive approaches and made considerable progress against NSF quality standards, aided by the increased use of incentive schemes for specific targets. Most impressively, the supportive and educative emphasis suggests a shift in culture

towards much more open sharing of information within and between organisations, all within a context of sector-wide organisational change and with limited financial resources.

Within secondary care, the early period saw an emphasis on leadership, accountability, reporting arrangements and baseline review for forward planning. While the need for an overhaul of the infrastructure of quality improvement systems was evident, it seems that the combination of short time-scale required for action planning and strong performance management resulted in pressures to rush the procedure, and in many cases led to a less thorough review than desired. The emphasis on structural changes for external quality assurance, evident in the measurement intentions of many directors with responsibility for implementing clinical governance and the lack of discernible strategic planning for organisational development, suggests that the quality improvement agenda seen by many as less of a concern.

A number of shared difficulties associated with the early implementation of clinical governance may be discerned, including unresolved tensions between quality assurance and continuous quality improvement. Such tensions require a delicate balance between formative and summative approaches, in order to avoid both clinical (medical) capture on one hand, and undermining the conditions of trust and honesty required for quality improvement on the other. While primary and secondary sectors faced similar difficulties, their differing contexts and traditions influenced early implementation activity.

In primary care, a tradition of individual general practitioners responsible for their own practice in the context of an emerging new corporate structure with relatively weak performance management mechanisms placed the emphasis on educative approaches to implementation. PCG/Ts began building a new corporacy through supportive quality improvement activity, further influencing behaviour through the use of financial incentives. Performance management functions were left largely to Health Authorities until their abolition. In contrast, secondary care implementation took place in the context of a relatively stable organisational environment, with a long history of mandatory quality improvement and assurance activity, with varying degrees of success. Implementation centred largely on structures and committee systems, driven by clear short-term goals from the centre. There is some evidence of greater emphasis on quality assurance mechanisms than building the longer-term quality improvement agenda. While primary care shows evidence of much educative activity, it is unclear how much of this activity focused on culture change as part of an integrated approach to organisational development. There is stronger evidence to suggest that in secondary care there was little co-ordinated activity explicitly designed to develop a culture of quality improvement.

The review identified a range of achievements in early implementation of clinical governance structures and processes. However, in their assessments of progress studies focused on different aspects of clinical governance, some concentrating on achievement against early development targets (Latham *et. al.* 2000), others looking at structural integration (Freeman *et. al.* 2001a) or barriers to implementation (Harrison & Lim 2000). Collectively, the reviewed studies shared two major limitations. Firstly, none used

validated measurement tools to assess progress in clinical governance at an organisational level. Such tools would allow an assessment of progress in clinical governance across the full range of clinical governance activities. Secondly, the reviewed studies concentrated on the views and activities of senior managers directly charged with implementing clinical governance, which may give a misleading picture of the effects of clinical governance activities upon the working lives of staff in health care organisations. Validated indicators of organisational progress and staff experience would be valuable for both internal and external audiences. Internal audiences would be able to identify areas of shortfall in order to plan their priorities for organisational development and to assess the outcomes of programmes designed to address those shortfalls; external audiences to assess progress across organisations.

CHAPTER 3

USING PERFORMANCE INDICATORS TO IMPROVE HEALTH CARE

QUALITY IN THE PUBLIC SECTOR: A REVIEW OF THE LITERATURE

"If performance indicators were living organisms the debate surrounding them would begin at conception and continue well beyond the grave. The controversy is understandable because during each phase of life, the performance indicators generate a clash of ideals between key stakeholders that can only be resolved through a compromise."
(Ibrahim 2001: p. 431)

"Performance management in the public sector... can be seen to exist as a rather inept 'science', often providing nothing more than crude indicators of an organisation's performance. ... These measures can be, and often are, set and monitored externally and used to hold organisations accountable for the way in which they use their allotted resources. Alternatively, such measures could be considered as 'indicators' of outcome. Analysis may then encompass a whole range of indicators, ... with targets now being set and monitored internally. These indicators can then be used in an investigatory mode, with priority placed on those that are of most significance to any particular organisation... In this situation, indicators may be seen as a management tool, used for internal control and development, rather than appearing as an externally applied source of accountability and implied threat."
(Boland & Fowler 2000: p. 418)

"Too much policing will not create a culture of quality. Rather it is likely to distort practice and prioritise the conventional and the measurable."
(Smith 1995: p. 308)

Introduction

Given the importance of performance indicators to current UK health policy outlined in chapter one, this chapter provides a systematic review of empirical and theoretical writings concerning their use to improve health care quality. The chapter outlines potential

problems and explores how best to derive, implement and use performance indicator data, presenting results thematically.

The two principal uses of indicator systems are as summative mechanisms for external accountability and verification, and as formative mechanisms for internal quality improvement. In the UK, the use of performance indicators in assurance and performance management systems has heavily influenced debate over their value, much of the criticism centring on their assumed objectivity. Major reported problems include potential perverse incentives; the validity and reliability of league tables derived from performance data; their potential to undermine the conditions required for quality improvement; and the difficulty of using such data to promote change. Technical problems include indicator selection; the availability, validity and reliability of data; confounding; and problems with robustness, sensitivity and specificity. Factors that help in the derivation, implementation and use of indicator systems include clear objectives, involvement of stakeholders in development, and use of 'soft data' to aid interpretation.

Health care is a complex adaptive system, in which staff anticipate the actions of those charged with controlling them. The use of performance indicators as the basis of reward or sanction is thus likely to lead to pressure to 'game' the system, leading to unintended negative consequences. While performance measurement systems may intuitively seem attractive, management of autonomous healthcare professionals requires a delicate balance of checking and trusting in order to deliver quality improvements.

3.1 The rise of performance indicators

Since the early 1980s, there has been an unprecedented increase in the number of health care performance indicators available in the UK NHS. Definitions of performance during this period have varied greatly, and at different points the emphasis has been on economy, efficiency, effectiveness, outputs, 'quality', outcomes, and social results (Talbot 2000). Consequently, the indicators themselves have also varied widely, in terms of what they purport to measure, their presentation and their intended audiences. Earlier initiatives, such as the Performance Indicators initiative of 1983, drew on activity and cost data for purposes of internal control by local managers. This emphasis shifted following the publication of *Working for Patients* in 1989, which gave new powers of regulation and performance management to the NHS Management Executive, and required indicators for the purpose of external accountability (Smee 2000). Thus, the increasing number of indicators at this point partly reflected changing trends in the governance of public organisations away from direct hierarchical control and towards indirect regulatory systems (Jacobs and Manzi 2000, McLaughlin *et. al.* 2001). Such 'hands-off' systems marry central control with local responsibility, and seek to verify compliance by auditing performance against targets (Power 1997; Carter *et. al.* 1992).

New Labour's modernisation agenda has again shifted the balance of public sector governance, further extending the role of indirect regulatory systems of audit (Newman 2000; Goddard *et. al.* 2000a). The White Paper *The new NHS: modern, dependable* (DoH

1997) outlined a new national Performance Assessment Framework (PAF) as the main instrument for securing management control in the NHS. The PAF marks an attempt to go beyond processes and outputs (Deeming 2001), and identifies six performance areas: health improvement; fair access; effective delivery; efficiency; patient/ carer experience; and health outcomes. Further indicators were announced with the publication of NHS Performance Ratings from September 2001, intended to provide a high level summary of overall performance of all NHS general hospitals, relating principally to patient experience. The trend to increased numbers of indicators shows no signs of slowing, as the number of organisations covered by, and range of indicators included in, the Performance Ratings is set to increase over time (DoH 2001).

This chapter summarises the general empirical and theoretical writings concerning the use of performance indicators to improve health care quality. Having outlined the aims and limitations of such systems, the chapter then identifies important lessons for their derivation, implementation and use in the context of health care quality improvement.

3.2 Methods

Data Sources, Selection and Data Extraction

The study comprised an extensive review of the electronic databases Medline, HMIC, ASSIA and BIDS selected for their depth of coverage of health care management topics. Searches were limited to articles published between January 1985 and May 2003 in peer

reviewed journals, using a simple truncated search for 'performance indica*'. All identified abstracts were reviewed, and full articles retrieved where it was clear that they contained empirical evidence or theoretical discussion on generic issues in the use of performance indicators in assessing health care quality. Where this was not clear from the abstract, full articles were retrieved before making a decision on inclusion or exclusion. Given the focus of the study, disease- and condition-specific papers without general import were excluded, as were papers detailing the reliability and validity testing of specific outcome indicators. Individual study methodologies were not evaluated given their largely discursive nature.

Data Synthesis

In total, 143 articles were selected, including reviews, original empirical studies and theoretical discussions. Given the review's focus on the broad themes of aims of performance indicator systems, perceived or reported problems, and perceived or reported factors facilitating their derivation, implementation or use, data extraction techniques were designed to facilitate the exploration of each theme and the inductive generation of sub-themes. Consequently, each article was carefully read to identify key findings under each theme, which were noted and recorded in a matrix (appendix 3). The matrix may be read either across (rows) to identify information relating to each of the themes within a specific article; or down (columns) to identify information relating to each of the articles on a specific theme. Sub-themes were generated inductively for each theme by reading the appropriate matrix column in sequence and summarising the major elements. Although this

is a comprehensive review findings are presented in summary form, illustrating important points.

3.3 Results

A careful reading of the literature reveals two broad traditions: the first is practice-oriented, prescriptive and optimistic of the value and use of performance indicators; the second is more critically engaged and sceptical. These literatures intersect at important points, and there is some agreement over what makes indicators desirable and productive and the difficulties involved in their production and use. The central ideal includes selection of appropriate indicators, notions of valid and reliable data collection, and sensitivity in the analysis and use of data in order to produce improvements in care quality. Both literatures are largely discursive, with an underdeveloped empirical base. Empirical contributions mainly comprise single- and multiple-site case studies of the introduction of performance indicator systems, and retrospective analyses of routine clinical data. The latter includes the adjustment and modelling of clinical outcome data as performance indicators.

Much of the practice-oriented literature focuses on matters internal to health care organisations and links conceptually to Total Quality Management (TQM) and Continuous Quality Improvement (CQI) approaches. Some writers make a significant contribution to the type of structures, processes and conditions required to develop performance indicator systems, yet this tradition tends to ignore analytic difficulties and treats the beneficence of

performance indicators as axiomatic. It does not consider conceptual problems, nor deal adequately with implementation difficulties. The academic literature focuses on the questions that the practical literature ignores, such as the unintended consequences of performance indicators, and the validity of inferences drawn from data. This literature is more consistently negative, and forms a discourse defined in terms of conceptual difficulties and statistical proof. These concerns are largely the result of the political context of the introduction of performance indicators in the UK public sector, specifically the use of performance league tables.

3.3.1 Aims of performance indicator systems

The two principal uses of indicator systems described in the literature relate to internal and external control and accountability (Davies and Lampel 1998; Rissel *et. al.* 1998; Smith 1995b) and formative quality improvement (McGlynn 1998; Goddard *et. al.* 1999). Boland and Fowler (2000) present a useful matrix structure of performance indicator systems based on two axes: the source of control (internal or external); and the nature of resultant action (positive [supportive and formative] or negative [punitive or summative]). Of the four possible ideal types, the authors contend that the practice of performance management in the UK public sector consists mainly of external, negative approaches. This contrasts with the use of performance indicators espoused by Continuous Quality Improvement (CQI) philosophies, which may be characterised as internal, positive approaches.

External, summative indicator systems act as mechanisms for external verification of quality improvements and central control (Smith 1995a; Smith 1990; Roberts 1994; Pollitt

1985; Martin and Kettner 1997; Shaw 1997). Such external accountability may be political, commercial or to the community (Leggat *et. al.* 1998), and is ultimately concerned with the renewal of legitimacy (McKevitt 1993). Information may be useful for purchasers and consumers (McKee and Hunter 1995; Popovich 1998), or help to inform political debate (Smith 1995a). Such systems are consistent with a mode of management concerned with verification of activity, which pushes control into organisational structures, inscribing them within auditable systems. By focusing institutional attention on their performance against indicators, governments seek to impose a policy agenda on organisations by embedding assumptions related to goals, values and purposes into the selection and structure of performance indicators (Jacobs and Manzi 2000).

Those championing performance indicators as a formative mechanism for quality improvement tend to do so from within a CQI paradigm, emphasising their potential to foster insight into practice (Portelli *et. al.* 1997). Here indicators are used as a focus for feedback and learning leading to improvement (Schyve 1995, and Tarr 1995). They act as a vehicle to align the objectives of staff and the organisation (Wyszewianski 1988), as stakeholders discuss and agree which indicators to include.

Carter *et. al.* (1992) classify indicators into three types: prescriptive; descriptive; and proscriptive. Prescriptive indicators are 'dials' that show achievement against targets, descriptive are 'tin openers' that record changes, and proscriptive indicators are 'alarm bells' that specify what should not happen. The two traditions each use indicators in exploratory and confirmatory fashion: acting as 'warning bells' for further investigation (Donabedian

1992; Holland *et. al.* 1994); or supporting existing suspicion and providing a trigger for action (Birch and Maynard 1986). Both traditions use benchmarking to identify poor performers and centres of excellence (Nutley and Smith 1998), making comparisons at the level of clinicians or units (Forster *et. al.* 1990). Yet, their information requirements and methods differ markedly (Table 3.3.1.1).

Table 3.3.1.1: Accountability and improvement approaches to indicator systems

	Assurance / Accountability	Improvement
Emphasis	Verification and assurance. Measurement oriented	Learning to promote continual improvement. Change oriented
Rationale	Provide external accountability and renew legitimacy	Promote change and improvement in care quality
Culture	Comparisons in order to make summative judgements on care quality. League tables. Blame and shame	Comparisons have a formative emphasis to learn from differences between providers and encourage improvement. Informal benchmarking to promote discussion and change
Precision required	High precision. Use of statistics to identify 'real' differences	Lower precision.
Epistemology	Empiricist. Statistical validity and reliability important	Interpretative. Use of other data sources and local information to provide context

Assurance systems tend to use indicators prescriptively, to make comparisons between providers leading to summative judgements on care quality; they seek to prove that providers with poor scores on indicators really are performing badly. This requires very high levels of data precision. The accuracy of information provided by performance indicators is assessed empirically, their claims to objective and material 'truth' settled with

recourse to statistical theory. The need for precision generates many problems, specified below. In contrast, indicators intended to provide information for quality improvement are used descriptively. They provide a provisional and partial truth interpreted in the light of local circumstance, and act as a starting point for discussion between stakeholders. While quality improvement models use indicators to develop discussion further, assurance models use them to foreclose it.

3.3.2 Perceived or reported problems with performance indicator systems

In the UK, the use of performance indicators in assurance and performance management systems has heavily influenced debate over their value. Many of the conceptual and technical problems below arise due to problems over validity, reliability and perverse incentives that tend to characterise such systems.

(a) Conceptual

In common with all assurance approaches, the primary goal of assurance focused performance indicator systems is the verification of improvements - quality must not only improve but be seen to do so. Yet, faith in measurement may be misplaced; in the context of external assurance, the existence of indicators does not remove the need for trust, but relocates it. Instead of relying on the internal control systems of professionals, auditors invest trust in data systems to provide information on practice (Power 1997). Under such circumstances performance indicator frameworks may simply displace existing informal modes of quality assurance. This is crucial, as because these strategies tend to be informal

and unrecorded, there is a risk of assuming that they do not take place. The irony is that new structures may displace these informal strategies, and in seeking to verify the accountability of agents may generate suspicion and fear, undermining the conditions of trust required for quality improvement (Sitkin and Roth 1993). The indicator framework may become a ritualistic verification system that does not provide reliable information on activity (Davies 1998). At its most ineffectual, a performance indicator system may simply provide information to fuel ritualised exchanges between agents and their principals. Deconstructionists characterise such exchanges as cultural performances and describe their publication- as league tables- as a form of contemporary spectacle or theatre (Stronach 1999).

The use of performance measures is clearly appropriate for monitoring compliance with regulations or comparing realities against a formal plan. However, indicators are of much less use in understanding how interventions or implementation processes have influenced results. They are not capable of showing why particular results are obtained, which is required in order to inform policy and programme modifications (Blalock 1999). While quality improvement oriented performance indicator systems may be able to use dialogue between stakeholders to generate insights, this avenue is difficult to pursue in assurance oriented systems.

(b) Technical

One of the main attractions of indicators is that they promise visible and concrete proof of performance. The argument is that indicators increase objectivity by applying commonly agreed rules of assessment across all organisations, thereby deriving knowledge that is independent of its creators. Yet critics question the very attribute of performance indicators that advocates of assurance based systems find so appealing; their objectivity. The claim to objectivity is essential to the use of indicators in performance league tables, yet poses many difficulties. Much of the debate is conducted in terms of statistical proof, but includes indicator selection, meaningfulness, robustness in the light of adjustments for confounding factors, as well as difficulties that arise when using them to inform service change. Indicators based on a limited range of items available in pre-existing information systems may have significant problems with their validity, reliability and comparability.

Indicator selection

Indicators provide information on a potentially limitless number of dimensions, and as indicator systems are unable to capture more than a fragment of what is important about the human experience of health care, some selection is required. A delicate balance needs to be struck between coverage and practicality: too few indicators and important aspects will be missed; too many and the instrument will be impractical to use and costly to maintain. By making some aspects visible, indicators marginalise other aspects and perspectives (Van Peursem *et. al.* 1995). They are thus conceptual technologies (Barnetson and Cutright 2000), shaping *which* issues are thought about and *how* people think about them.

The structured reporting of indicators further obscures the complexity, contextuality and layered meanings involved in interpretation. Variation in definitions of what appear to be 'simple' indicators at different centres may lead to the failure to compare like with like (Jackson 2001; Gross *et. al.* 2000; McColl *et. al.* 2000). Further difficulties in operationalising indicators arise specifically in the public sector. This is partly due to the lack of a financial 'bottom-line', but also because of the existence of multiple conflicting objectives, the over-riding importance of political objectives, and short political time-horizons (Hepworth 1988). There are no technical solutions to these problems, and value judgements are required given the existence of legitimate political debates surrounding the definition of 'appropriate' measures, such as whether to take social factors into account (Stewart and Walsh 1994).

Data: availability and reliability

There are often problems with the availability of data, resulting in a tendency to focus on measuring what there is data for, rather than items that correlate with the system's goals and objectives (Lorence *et. al.* 2002; Lowry 1988; McKee and James 1997). Data collected by clinicians has the benefit of encouraging discussion and debate for health improvement (Rowan and Black 2001), but at the potential cost of making external benchmarking difficult (Nyhan and Marlowe 1995). Data accuracy is also important, the issue being whether between group differences reflect quality of data rather than quality of care (Kazandjian *et. al.* 1996). Flawed data is likely to increase confrontation and reduce co-operation (McKee and Hunter 1995), especially when used for summative assurance purposes. All data collection relies on the goodwill of clinicians, and is thus susceptible to

manipulation, particularly when reward and censure depend on results (Audit Commission 2000a; Diamond *et. al.* 2001).

Data: validity and confounding

Even where data is available and reliable, it may be potentially misleading and easily misinterpreted (McColl *et. al.* 1998; Smith 1995b). Measurement validity reflects the extent to which indicators truly represent a more abstract variable. In order to be valid measures of health care quality, indicators need to reflect attributes of the health care system, rather than attributes of the patient or of other non-healthcare characteristics. For example, readmission rates are a valid indicator of care quality to the extent that readmission is due solely to deficiencies in the quality of the previous care. Because measures, and especially outcomes, are affected by other variables, it is difficult to attribute variation to performance (Leyland and Boddy 1997; Bull *et. al.* 1994; Leng *et. al.* 1999; Milne and Clarke 1990; Orchard 1994). The meaning of indicators thus becomes debatable. The central point is that indicators should only relate to factors that are under the control of those under scrutiny (Parry *et. al.* 1998; Hauck *et. al.* 2003).

To avoid confounding, all other exogenous and endogenous factors affecting the indicators must be controlled (Birch and Maynard 1986; Barnsley *et. al.* 1996; Davenport *et. al.* 1996). Potential confounding factors include configuration of the local health economy (Brown *et. al.* 1995; Carter 1989), socio-economic variations (Giuffrida *et. al.* 1999a), case-mix, co-morbidity and severity (Mant and Hicks 1996; Rigby *et. al.* 2001). Without adjustment, it is not clear to what extent indicators identify the contribution of health

services to health care (Mulligan *et. al.* 2000). There are particular problems associated with outcome measures, as they may occur over long time scales and suffer from problems of attribution, especially in measures of chronic illness (Smith 1995c). Theoretical models incorporating antecedent factors have been identified as a solution (Coyle and Battles 1999; Ashton and Wray 1996), but have data requirements that are unlikely to be provided by routine systems (Challiner *et. al.* 2003), in excess of that required to explore processes (Mant 2001). Data collection costs also need to be considered, as organisations may face significant financial and time costs in producing and reporting the data, potentially incurring significant opportunity costs (Heasel 2000; Boyne 2000).

Dealing with confounding

Risk adjustment models may control the effects of confounding. Four principal methods are available, each posing its own difficulties (Blumberg 1986; Giuffrida *et. al.* 2000). The first is standardisation, in which differences in variables such as age and sex mix in a population are standardised. The problem is that there may be other population characteristics that may confound the indicators, and we cannot standardise for them unless we know what they are and hold data on them. The second is cluster analysis, which groups units for comparison into clusters sharing similar socio-economic profiles. The problem is that clustering criteria are arbitrary. The same clusters are often used for all measures, which is plausible only if the variables used in the clustering procedure affect all indicators in the same way. Options three and four are both ways of modelling data. Data Envelope Analysis (DEA) (Giuffrida and Gravelle 2001) calculates lowest possible indicator rate with given levels of confounding variables, and compares an areas actual

performance against the best possible rate given its confounding profile. The final option, multiple regression analysis, uses information from all units to predict the indicator rate that a unit should have given the values of its confounding variables. The problem with modelling the data is that residuals that cannot be explained by the model are due to other factors- these may include care quality, but could also be due to any other potential variable not in the model. Further, poor practice is included in the norming procedure (Leyland *et. al.* 1991). Perhaps the greatest difficulty is that the two final options also suffer from not being transparent to the end user. The question becomes 'do you trust the statistician', as attempts to increase the validity of the data may simply undermine its credibility. The tendency of different adjustment methods to yield different results brings the robustness of the process further into question (Brown *et. al.* 1995; Nutley and Smith 1998).

Indicators: robustness, sensitivity and specificity

Indicators may falsely convey an impression of objectivity to what is often weak and ambiguous evidence (Davies and Lampel 1998; Van Peursem *et. al.* 1995). Small numbers of cases mean low significance (McGlynn 1998; Sheldon 1998), the data requirements for precision are excessive (Mant 1995; Logan 1991), and random variation in measures may be misinterpreted (Smith 1995c). The inability to measure social outcomes with precision creates problems for accountability, as comparisons become difficult and potentially deceptive. Year to year variation within league table rankings has been high and unreliable, and consequently league tables may lead to unnecessary praise or sanction (Wellard 2000). It is important that indicators are able to identify all poorly performing units (sensitivity), and that all units identified by indicators as performing poorly really are performing poorly

(specificity). Poor sensitivity and specificity results in false assurance or denigration, where indicators incorrectly identify individuals / organisations as poor, or fail to identify the poorly performing (Davies 1998; Forster *et. al.* 1990; Goldstein and Spiegelhalter 1996). It should be noted that different levels of sensitivity / specificity are needed depending upon purpose, with summative accountability decisions requiring much higher accuracy than formative developmental ones. Given the disputed nature of their meaning and precision, prescriptive indicators are often used descriptively in practice, to provide a starting point for discussion (Carter *et. al.* 1992; Linder 1991). Indeed, the need to take context and other informal 'soft data' into account is widely recognised (Goddard *et. al.* 1999; DoH 2001).

Indicators: promoting action and change

Indicators have an end-of-process focus, which takes time to collect data and act (McColl *et. al.* 1998b; Davies and Lampel 1998). They are also a poor motivational device for action and change (Mannion and Goddard 2001). All indicator systems give rise to perverse incentives and unintended consequences (Smith 1990; Smith 1995a; Goddard *et. al.* 2000b; Roberts 1994; Greener 2003). This problem is due to the assumption of a mechanistic response to performance indicator feedback, in which provision of information leads directly to changes in practice. Health care may be better characterised as a complex adaptive system (Plsek *et. al.* 2001), in which people anticipate the reactions of those charged with controlling them. Accordingly, the behaviour of those within a system may give rise to negative unintended consequences (Thompson and Lally 2000). Smith (1995a) outlines the potential distortions induced by over reliance on performance indicators by

drawing attention to the implicit management incentives of such schemes. Drawing on a wealth of literature, he concludes that the almost universal finding is that performance indicators distort behaviour in unintended ways (Table 3.2).

Table 3.3.2.1: Unintended consequences of public sector performance indicator systems (After Smith 1995)

Tunnel vision	Emphasis on phenomena quantified in the measurement scheme
Sub-optimisation	Pursuit of narrow local objectives, rather than those of the organisation
Myopia	Pursuit of short term targets
Measure-fixation	Pursuit of strategies enhancing the measure rather than the associated objective
Misrepresentation	Deliberate manipulation of data
Misinterpretation	Drawing misleading inferences from raw performance data
Gaming	Deliberate manipulation of behaviour to secure strategic advantage
Ossification	Organisational paralysis due to rigid performance evaluation

When rewards are dependent on data held by professionals, dysfunctional behaviour may result. Potential effects include manipulation of records (McKee and Hunter 1995), and even changes to clinical practice itself (Smith 1995b), potentially reducing the number of surgical interventions used on high-risk cases (McKee 1997b). Hannan *et. al.* (1994) dispute this, in a secondary analysis of cardiac surgery outcomes showing no such reduction in surgery in high-risk cases.

A further problem with the use of outcome 'rates' as indicators is that they provide no indication of what, if any, action is appropriate to improve healthcare. They conceal the

detail required to show what went wrong (McColl *et. al.* 1998b). The problem is that even if we can make meaningful comparisons and identify deficiencies, it may still be unclear what action to take (McLaughlin *et. al.* 2001), which requires CQI based interventions with clinical teams to improve clinical quality.

3.3.3 Perceived or reported factors facilitating the derivation, implementation and use of performance indicators

Much of the literature on health care performance indicators is concerned with the technical and conceptual difficulties involved. This is unsurprising, given that those developing indicators rather than those charged with using them most often write it. However, the big question for practitioners is how to use indicator data most effectively to encourage and develop quality improvement. Those with responsibility for implementing indicator systems need to bear in mind the need for an analysis of the work environment; design of indicators; integration into the workplace; and evaluation and further development of indicators (Cave *et. al.* 1990). While it is rare to conduct an initial analysis of the work environment, it is also critical as it affects the ability of the various interest groups to engage fully with the process. The existence of staff hierarchies or a history of poor experience with quality improvement initiatives will affect the level of staff involvement and support of any initiative. Those charged with implementing work on performance indicators would be wise to identify and address such barriers before developing indicators.

Derivation

It is vitally important that stakeholders share a common understanding of the intended use of proposed indicators. Internal quality improvement and external accountability are distinct requirements, and the information and processes required for each differ markedly (Solberg *et. al.* 1997). External assurance purposes require investment in reliable and valid data collection, and modelling to adjust for confounding. Quality improvement approaches require less robust data but require processes that encourage discussion of results and lead to service improvement. A lack of clarity over the aims of an indicator system will inevitably lead to problems over ownership of the data and disputes over its meaning and proper use. It is therefore important that all participants agree on the use of indicator data, and design systems appropriate for the specified task. Before indicators are developed, stakeholders should clearly define their goals (Smith and Frowen 1997; Clarkson and Challis 2002). All principal stakeholders need to be involved in the development of indicators (Nyhan and Marlowe 1995; Smith 1995a; Goddard *et. al.* 2000a). This includes clinical, managerial, purchaser and patient perspectives (Popovich 1998). A number of indicators are required to ensure all dimensions of the area of interest are covered. The need for multiple measures should needs to be balanced against the selection of only those indicators critical to the agreed goals (Leggat *et. al.* 1998). Finally, indicators are only meaningful if they are markers of outcomes or processes that are under the influence of clinicians (McKenzie and Shilling1998). If clinicians are unable to affect the result, the indicator is redundant.

Implementation

Well-derived indicators can act as a catalyst for change within the organisation (Tarr 1995). While a sensitive development will minimise resistance, managers do need to anticipate resistance during the implementation process (Meekings 1995). Executives may signal the importance of the initiative by active 'championing' and releasing adequate resources for its successful implementation and delivery (Holloway *et. al.* 1999; Tan 1999). While executives can point the way, line managers have an important role to play in ensuring top-to-bottom links in the organisation (Nyhan and Marlowe 1995). By using indicators in a transformational manner, managers and clinicians may integrate quality initiatives around an agreed agenda. Indicators should not be used simply as an additional extra, but integrated into everyday working practices, as part of both clinical and management processes, and the wider organisational culture (Flynn 1986).

Implementation requires the development of IT systems required for information capture. Wherever possible, the automated collection, input, analysis and retrieval / dissemination of data is helpful (Al-Assaf 1996). The physical provision of IT systems is of course only the first step, as data quality (Leggat *et. al.* 1998) and reliability (Schyve 1995) require assurance processes.

Use

Given the difficulties involved in the validity of indicators, there is a fine line to be drawn between the use of unadjusted data and attempting to control for confounding factors. Too

much adjustment and the immediacy and credibility of indicators is lost; too little adjustment and indicators may be very misleading. Where possible, it is helpful to adjust raw figures for social and environmental factors, plus severity and case-mix (Holland *et. al.* 1994; Leyland *et. al.* 1991). However, users should remember that there are no technical solutions to problems of interpretation. Reported indicators are exactly that; indicators that focus attention on issues of interest. They are neither proof of a problem or its solution. Users would be well advised to offset precision and reliability issues with a holistic impression derived from a wide range of indicators, and link interpretations to other local sources of knowledge (Smith 1995a; Goddard *et. al.* 1999). From this perspective data are sensitising material only, providing hints to outcomes emerging from activity and subject to debate and discussion (Smith 1995c; Krivenko and Chodroff 1994). Much of the debate around statistical significance of results can be avoided, as adverse occurrences in small numbers of cases require individual case analysis to explore circumstances, not aggregation into population rates (Logan 1991).

In order to be consistent with the above, participants require learning focused, non-judgmental feedback (Greene 1999; Nyhan and Marlowe 1995; and Tarr 1995). Indicators should be used to learn and correct rather than blame, with room for discussion and management judgement. Ideally, multi-disciplinary 'performance improvement teams' should discuss identified problems (Kazandjian *et. al.* 1996) and agreed planned changes in the light of their deliberations. In order to develop co-operation, users would be wise to avoid external release of indicators designed for internal quality improvement purposes. The control of data should rest with those responsible for service delivery and

improvement, to avoid undermining the trust necessary for quality improvement. Finally, goals, indicators and their underlying values require constant re-evaluation in order to ensure continued relevance to the organisation (Portelli *et. al.* 1997).

Conclusions

The literature describes two principal uses of indicator systems: as a summative mechanism for external accountability and verification in assurance systems; and as a formative mechanism for internal quality improvement. The weight of evidence considered in this review suggests that the use of performance indicators in a summative way as a basis for praise or sanction is almost inevitably corrosive and corrupting of the indicators themselves. Such accounting systems place trust in systems rather than individuals, further undermining the conditions of trust required for quality improvement. A range of technical problems arise due to the precision of data required to make summative comparisons, and further negative unintended consequences follow due to the pressures on clinicians and managers to 'get good results'. Of far greater potential benefit is the formative use of indicators as clues to performance, discussed and interpreted by clinicians and managers in the light of local contexts and with the aim of continuously improving the quality of clinical care. Such approaches foster trust and communication between clinicians and managers, with the result that they are better able to work through problems with care delivery and improve quality.

It is clear that indicators of health care quality are not axiomatically good. They may be inaccurate, misleading and even dangerous in their negative unintended consequences. They are very seductive as they promise an objective view of health care quality; yet, it is a promise that they are ultimately unable to keep. Those using indicators of health care quality should do so with caution, and remember that they are precisely that: indicators, for further investigation and discussion, requiring cautious interpretation in the light of local circumstances. It is encouraging that documentation outlining the PAF explicitly accepts the need for such cautious interpretation. Failure to do so in practice will generate strong perverse incentives, and the pressure on clinicians and managers to 'game' the system will be very difficult to resist (Hargreaves 2003).

There is a danger of overplaying both the objective nature of summary measures of health care quality, and the usefulness of centrally collated data in encouraging improvement. The intention behind the NHS Performance Ratings is to gain an overall view of the performance of trusts, by aggregating indicators. The system extends the use of performance indicators beyond accountability and into future resource allocation. Classification will have direct effects on organisations; those fortunate enough to be 'three star' will have considerable freedom to develop local services; others will have less freedom, and receive scrutiny and developmental advice from the Modernisation Agency.

The danger is that we may be embarking on another round in the 'natural cycle' of assurance based indicator systems (Cave *et. al.* 1990). The cycle starts when a crisis flags the need for accountability. This triggers new indicators and a drive for improvement. To

maintain improvement, managers have an incentive to manipulate ('game') the indicators.

An attempt to correct this via a proliferation of indicators leads to a complex system that is difficult to understand then falls into misuse as attention shifts within the organisation.

There is potentially much to be gained by the sensitive and cautious use of performance indicators, and much to be lost if they are used without due caution. Critics of performance measurement are not questioning the need for measurement and accountability; what is at issue is the most effective way of improving health care quality, and the role for performance indicators within the overall strategy. Thus, the underlying philosophy of the use of the data is all-important. Wherever data is used to judge performance, professionals expend energy in beating the system. They become adept at grabbing the carrots and avoiding the sticks, and the original aim of improving quality becomes lost (Goddard *et. al.* 1999). However, if staff trust that data will be used to reflect on practice and that their motives are not in doubt, they will be more likely to collect and use data.

Such a hands-off approach does not imply an abdication of responsibility; rather, it suggests the need for a delicate balance between checking and trusting (Davies and Mannion 1999a). While managers need to be able to assure the public of care quality via PAF and 'stars' indicators, practitioners need time and space in an atmosphere free from threat to identify and exploit opportunities for improvement. The tension between assurance and improvement lies at the heart of clinical governance, whose long-term cultural agenda requires an emphasis on quality improvement.

CHAPTER 4

MEASURING PROGRESS IN CLINICAL GOVERNANCE: CURRENT OPTIONS AND THE CASE FOR ADDITIONAL MEASURES

"... heavy-handed external assurance may destroy the conditions of trust, honesty and innovation that are required for successful internal quality improvement."
(Davies & Mannion 1999a: p. 10)

Introduction

As outlined in chapter one, the assurance and continuous quality improvement paradox lies at the heart of clinical governance (Freeman *et. al.* 2001) and from its inception they have been uneasy bedfellows. Quality measurement can be undertaken in either a summative, judgmental way or in a formative and supportive manner. The NHS Plan (2000) constitutes an attempt to weld these two together under a system of earned autonomy:

"Thus we see policy approaches apparently striving to balance checking and trusting: aiming to get measured evidence of performance and encouragement for improvement without stifling intrinsic motivation and local initiative." (Davies & Mannion 2002: p.60)

This chapter explores available measurement options for assessing progress in clinical governance, deploying the conceptual framework of Boland & Fowler (2000) in order to categorise approaches to measurement on two axes according to the source of control (internal / external) and the nature of the resultant action (summative / formative). After

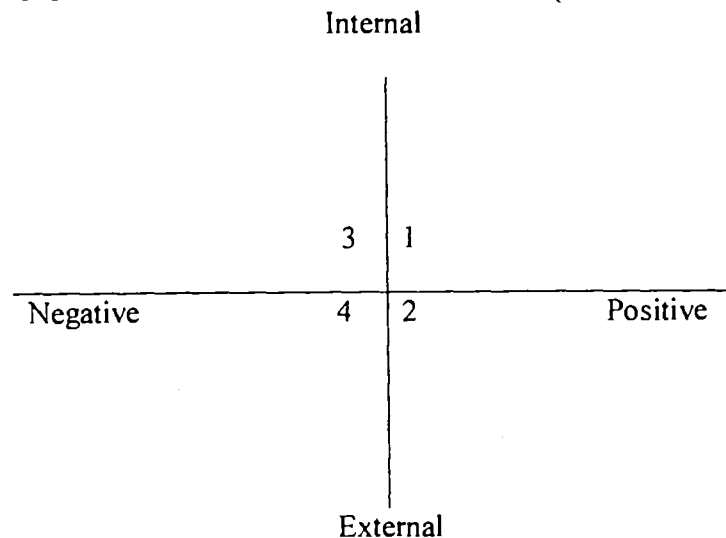
reviewing the current options and making the case for additional internal and formative measures to support the quality improvement agenda, specific cases are made for two measures. The first is for a measure of clinical governance climate based on staff perceptions. Conceptual and methodological difficulties in the pursuit of measures of organisational climate are identified and explored, specifically the notion of climactic homogeneity and issues concerning the aggregation of results. The second measure is of organisational competence in clinical governance, measured at board and directorate manager level and based on variance between respondents.

4.1 Measurement options: a conceptual framework

There is a wide and potentially bewildering array of measurement indices for assessing the quality of health care provision at the level of local health economy, organisation, clinical directorate or service team. Options include accreditation, quality management frameworks, performance management systems, benchmarking, and statistical process control, among others. As outlined in chapter 1, within the UK the role of central government in health care quality improvement is increasingly prescriptive, consisting of a broad range of new quality systems operating within a performance management context (Ferlie & Shortell 2001). New systems include an expanded Performance Assessment Framework; visits, inspections and latterly performance ratings ('stars') from a Commission for Health care Audit and Inspection (CHAI); a Modernisation Agency to support organisational process redesign; and national quality targets, such as shorter waiting times for appointments. New structures of clinical governance within health care provider units are designed to ensure local delivery of the national agenda.

Given this level of complexity, the need for additional measures of progress in clinical governance may legitimately be questioned. Yet, the case for additional measures may become clear if measurement options are considered within a conceptual framework of performance measurement, performance indicators and improvement initiatives as applied in public sector organisations. Figure 4.1.1 illustrates such a framework (Boland & Fowler 2000), consisting of a matrix structure based on two axes: the source of control (internal or external); and the nature of resultant action (positive [supportive and formative] or negative [punitive or summative]).

Fig 4.1.1: Control locations and resultant action matrix (Boland & Fowler 2000)



The vertical axis identifies the source of authority for control, either from within the organisation (internal) or from outside (external). While the former implies using

indicators for internal purposes, the latter implies the existence of an outside body holding the organisation to account for its actions. The horizontal axis concerns perceptions of the nature of resultant controlling actions: supportive and formative (positive) or punitive and summative (negative). Negative action implies an assumption that under-performance is due to inefficiency, leading to actions such as a reduction in resource allocation. In contrast, positive action implies that in the same situation an investigation would be undertaken to explore why the situation had arisen, leading to redeployment of resources or organisational development activities such as staff development. Considered as two independent dimensions of control location and resultant action, the axes combine to produce four potential options:

1. Quadrant 1: Internal control, positive action

These typically include internal quality assurance procedures, leading to supportive actions. This approach is broadly characterised by Total Quality Management approaches, benchmarking activity and statistical process control. Measurements are taken in order to support internal quality improvement.

2. Quadrant 2: External control, positive action

This quadrant includes external bodies charged with auditing the performance measurement and quality improvement systems of organisations i.e. audit of systems at one step removed, in order to offer advice on how the whole organisation can develop component parts to ensure continued improvements. It is positive in nature given its supportive, largely nurturing role.

3. Quadrant 3: Internal control, negative action

This concerns measurement of internal sub-units as in quadrant 1, but here negative actions may arise. Although data may be gathered internally, in practice information is often required by an external agency. Thus the collection and reporting of waiting time data in hospitals is an internal concern, but also required by a central agency. As explored in chapter 1, new public management approaches rely heavily on the local collection of data used externally to monitor performance.

4. Quadrant 4: External control, negative action.

As in quadrant 3 above, organisations may expect to be judged on their activities with punitive consequences for poor performers. However, in this quadrant authority is imposed through lines of accountability to an external agency.

Boland and Fowler (2000) contend that while quadrant 1 is the most desirable location in terms of satisfying the long-term needs of the majority of stakeholders, the practice of performance management in the public sector lies ultimately in quadrant 4 given its concern with accountability. They specifically identify clinical governance as an example of external control with negative action. This may be overstating the case, as the combination of internal quality improvement and external quality assurance at the heart of clinical governance means that activities may be discerned within all four quadrants (Fig 4.1.2). However, the balance of performance mechanisms associated with clinical governance seems firmly within quadrant four, external control and negative action.

4.2 External approaches

External approaches are in essence regulatory. Walshe (2002) identifies the key features of regulation as the involvement of a third party- the regulator- in market transactions and interorganisational relationships, placing responsibility for overseeing performance with an entity- the regulator. He identifies two 'paradigms of regulation' to which regulators typically subscribe, deterrence and compliance. The former assumes that those regulated need to be forced to behave well through strict enforcement of demanding standards, lest they pursue their own interests to the detriment of all others. They require searching judgement and it is the duty of the regulator to bring them to account. In contrast, the compliance model assumes that those regulated are well meaning and seeking to comply, requiring support and advice from their external assessors. The deterrence and compliance models equate respectively to the 'summative' and 'formative' approaches outlined by Boland and Fowler (2000), and may thus be incorporated into the quadrant framework (Figure 4.1.2)

A review of the scope, mechanisms and use of external peer review techniques (ExPeRT) operative in western European health care systems (Shaw 2000) identified four assessment mechanisms: accreditation; ISO 9000; EFQM; and *visitatie*, each of which are used in the UK (Heaton 2000). Klazinga (2000) embeds these assessment mechanisms within broader perspectives of external review: organisational (accreditation), process control (ISO 9000), quality management (European Foundation for Quality Monitoring Excellence Model EFQM EM), and professional (*visitatie*). It should be noted that not all of these assessment

mechanisms are summative; both *visitatie* and the EFQM EM are formative in intent, and may be located in quadrant 2 of the action matrix (figure 4.2.1). These two have also been explicitly recommended or incorporated into measurement mechanisms for clinical governance (Freeman 2000; Jackson 2000).

Fig 4.1.2: Control locations and resultant action matrix in the UK NHS

		Internal	
Waiting time targets		Statistical process control	
		Global clinical governance measures CHI self-assessment	
	3	1	
Negative	4	2	Positive
Accreditation (organisational)		<i>Visitatie</i> (professional)	
ISO 9000 (quality systems)		EFQM (quality management)	
Commission for Health care Audit and Inspection (CHAI)		Benchmarking	
Performance Assessment Framework		Modernisation Agency	
Performance Ratings ('Stars')			
National Service Frameworks			
		External	

Traditionally, the focus of the standards in each model reflected their original purpose: professional standards (*visitatie*), health service delivery (accreditation); management systems (EFQM); and quality assurance systems (ISO). Models are typically complementary rather than mutually exclusive, the differences in uptake in any health care

system being a function of the attributes of the models and the interests of different professional groupings (Heaton 2000). Thus, debates over models and methods are ultimately concerned with professional control, and the perceived value of approaches is dependent upon judgements concerning the appropriate balance between professional self-regulation and accountability. At the level of health care organisations clinical professionals tend to be supportive of visitatie and accreditation. Their enthusiasm is however dependent on who is going to use the model and to what aims, and they will seek to limit accreditation programmes to areas of professional performance that they deem appropriate. Similarly, managers may either promote or oppose visitatie depending on their perspective on the value of self-regulation, and those managers intent on increasing managerial control will tend to pursue accreditation and ISO. At the government level, debate amongst policy makers ultimately concerns the ability of the mix of models to achieve a balance of power between stakeholders congruent with government policy towards accountability (Klazinga 2000). Thus future development in external peer evaluation is primarily dependent upon the way that responsibilities for quality assurance are assigned.

As outlined in chapter one, the recent exponential growth in external regulatory mechanisms has been occasioned by increased decentralisation and use of regulated markets in health care service delivery, with consequential greater emphasis on oversight methods of accountability (Power 1997; Hood *et. al.* 1998). Consequentially, their use strikes new balances of trust and control between patients, professionals and government:

"In essence, external quality assurance mechanisms ... are the operationalization of control mechanisms within health care systems. The hard edges of the term 'control' can be mitigated by speaking about mechanisms for accountability or mutual trust, but the bottom line is that they should not only assure, or even better improve, the quality of an assessed service, but should also assure, or even better improve, the trust of external parties such as patients, financiers and government." (Klazinga 2000, p.455)

Consistent with the growing influence of managerialism, since the early 1990s the European Union experience is of convergence between external peer review models, each moving towards comprehensive standards for organisation, management and clinical performance. Thus, accreditation systems are embracing ISO 9000 standards (Heaton 2000) and EFQM is being used as a framework within which the other models are incorporated (Moulin 2002). Notwithstanding such pressures of convergence, the balance of formative and summative elements within these various approaches to external review remains. Further details on each of the identified mechanisms in Fig 4.1.2 appear below.

4.2.1 External summative approaches

4.2.1.1 Health care organisation perspective: Accreditation

Organisational accreditation originated in the USA with Joint Commission on the Accreditation of Healthcare Organisations, and was exported to Canada via Australia and then to Europe in the 1980s (Shaw 2000). Usually voluntary in nature, accreditation is typically undertaken by a multidisciplinary team of health care professionals against published standards. In the UK, accreditation is becoming increasingly widespread, evinced by a wide range of programmes such as the Hospital Accreditation Programme (HAP), the Kings Fund Organisational Audit (now HQS), Investors In People (IIP) for

training and development of staff, and Clinical Pathology Accreditation (CPA) (Heaton 2000).

4.2.1.2 Quality systems and process control perspective: ISO 9000

International Organisation for Standardisation (ISO) standards for quality systems consist of a series of generic statements, conceptualising an organisation largely in process terms. This is justified by the fact that the generic norms are applicable to all types of organisation and certification relates to the quality system – not the content of the work:

“Applied to health care, it (ISO 9000) helps to strengthen the process orientation of a health care service but will not give any assurance about the appropriateness of the selection of treatments for patients or the health outcomes.”
(Klazinga 2000).

ISO standards have been used to assess whole health care organisations, but as the standards principally relate to administrative procedures they are most often used in mechanical departments such as laboratories (Shaw 2000).

4.2.1.3 Commission for Health Improvement (CHI) and Commission for Health care Audit and Inspection (CHAI)

CHI was created under the 1999 Health Act as an independent body charged with monitoring performance and facilitating improvements in clinical care quality, through routine reviews and special investigations. This constitutes a potential double paradox: that of supportive and developmental in the context of reviews and the hard-edged investigator charged with naming and shaming; and the extent of its independence from government,

given that Secretary of State may request investigations. The dual identity has been discussed by classical allusion to Janus, the roman god of beginnings with two faces (Rowland 2003); and more prosaically by posing the question CHI: quality police or a midwife of change? (Day & Klein 2001).

Since its inception, senior CHI personnel have consistently sought to place the emphasis firmly on facilitating improvement. While generally well received within the service, Day & Klein (2002) raise a number of important tensions within CHI clinical governance reviews. These include the combination of summative and formative dimensions of the reviews; confusion over the balance between checking for the implementation of mechanisms or assessing their impact; queries over data quality and the extent of critical evaluation of data in context; and all within highly heterogeneous organisations.

On the whole CHI seem to have managed the paradox of investigator and supporter, and the more developmental and enabling face has tended to prevail (Rowland 2003). However, its dual roles of assurance and improvement have provoked ambivalence and suspicion on the part of those on the receiving end of reviews, and some have reported negative experiences:

"The original philosophy behind CHI, strongly espoused by its most senior staff, is all about development and learning, motives and methods that are not only worthy, but likely to be effective. Are we more likely to be truthful when we talk to the nurse with our interests at heart, or to the nasty customs officer who is just out to police the excise laws? But at least we know where we stand with the customs officer. He wears his nasty official's uniform, and makes no pretence to be our friend. It is the disingenuousness of the CHI process that not only hurts so much, but also ensures that it will ultimately not achieve its aims. The trust was seduced and betrayed, and having been deceived will be more wary next time." (Shapiro 2003: p. 18)

These pressures are likely to be exacerbated with the development of CHI into the Commission for Health care Audit and Inspection (CHAI). The NHS Reform and Healthcare Professions Act (2002) and Delivering the NHS Plan (2002) outlined new powers and responsibilities for CHI, along with the new name. Its expanded portfolio includes performance management, waiting lists and value for money. These include the creation of an Office for Information on Health Care Performance, licensing private health care provision, conducting value for money audits, validating published performance statistics, publishing star ratings and an ability to recommend special measures.

The CHI model is thus becoming increasingly summative. Ministers increasingly see CHI as 'quality police' (Day & Klein 2001), yet this appears to clash with CHI as a supportive developmental agency, working in partnership with NHS organisations. Time will tell whether the desire to be supportive and developmental will be able to survive the involvement in performance management, waiting lists and value for money work. The resignation in April 2003 of CHAI's CEO, who had previously held the same office at CHI and was known for his commitment to formative approaches, heralds greater uncertainty.

4.2.1.4 Performance management: NHS Performance Assessment Framework

Central influence over policy implementation in England has increased dramatically under reforms instituted by New Labour, principally via Public Service Agreements between Whitehall departments and the Treasury, and service delivery agreements explaining how

the high-level targets contained in the public service agreements are to be reached (Lee & Woodward 2002). In the context of health care, this agreement is implemented through the Performance Assessment Framework (PAF); it is the routine quantitative instrument for monitoring the progress of health care organisations under the 1997 reforms (NHSE 1999). Six areas of the new PAF were agreed in 1999: health improvement, deaths from all causes; fair access; effective delivery of appropriate health care; efficiency; patient / carer experience of the NHS; and health outcomes.

4.2.1.5 Performance Ratings ('stars')

Three sets of information are combined to produce a summary measure of performance from three star (excellent) to zero star (poor). The first area of information is performance under key targets, such as waiting times. The second comprise other measures under the headings of patient, clinical and capacity focus. Thirdly, where available CHI review information is used to adjust the star rating. Star ratings have an important effect on the earned autonomy of trusts, as only three star trusts are allowed to spend improvement fund monies at their discretion or apply for foundation status. Crucially, trusts failing badly on the key targets will receive no stars, no matter how good the CHI report or how well it performs under the other measures. Given the possibility of reward or sanction, the potential for perverse incentives leading to misrepresentation is very high.

4.2.1.6 National Service Frameworks (NSFs)

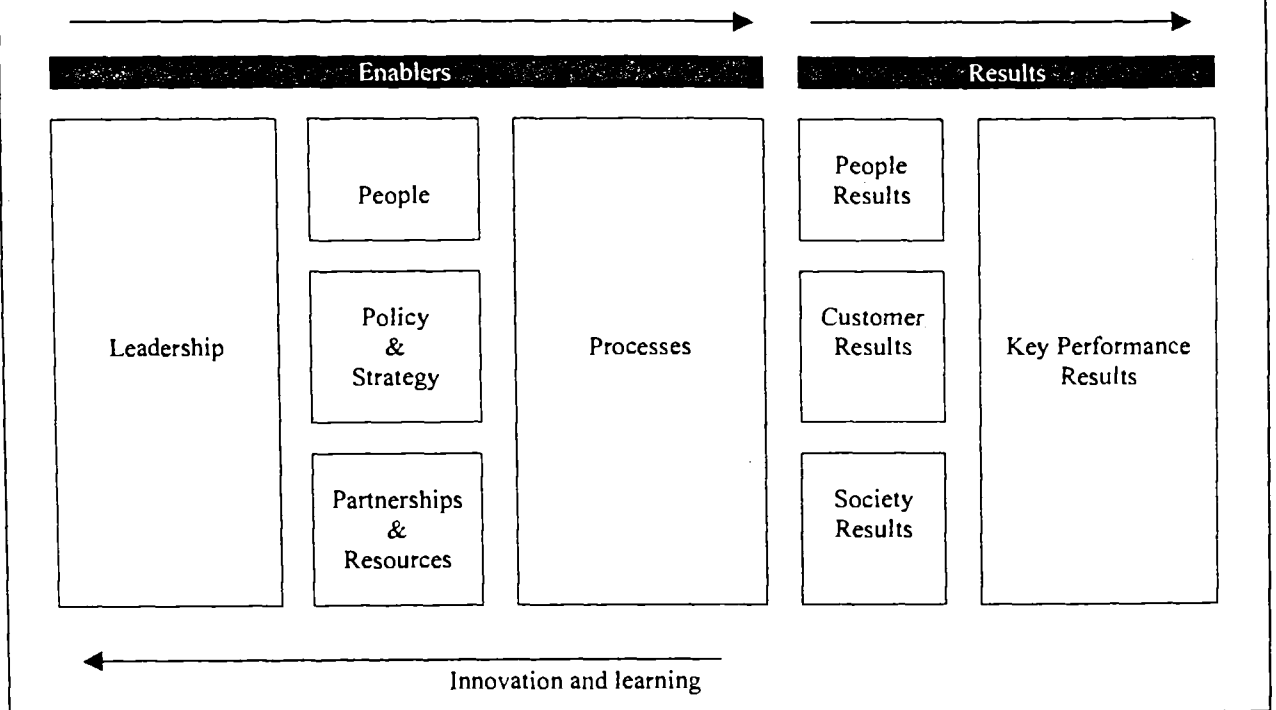
The intention behind NSFs is to set national standards and define service models for specific services or care groups including, mental health (DoH 1999), coronary heart disease (DoH 2000), diabetes (DoH 2003) and older people (2001); and establishes implementation programmes and performance measures against which progress may be measured. Each NSF includes national standards of care which patients' may expect, a definition of the evidence base associated with interventions and their associated costs, as well as work commissioned to support implementation such as research, appraisal, benchmarks and outcome indicators, and supporting programmes including workforce planning, education and information management.

4.2.2 External formative approaches

4.2.2.1 Quality management development perspective: EFQM EM

In essence, the EFQM Excellence Model subscribes to Deming's continuous improvement philosophy of 'plan-do-check-act', and is results-oriented, driven by a process of self-assessment. It consists of nine criteria (figure 4.2.2.1.1), of which five relate to how things are done ('enablers'), and four to what is achieved ('results'). Each criterion is divided further into sub-criteria defining excellent practice. For example, leadership sub-criteria include the development of a mission, vision and values and ability of leaders to act as role models of excellence (EFQM 1999).

Figure 4.2.2.1.1: the EFQM Excellence Model



© EFQM (1999)

The model may be used to measure success and set improvement targets via self – assessment, using teams drawn from service areas, possibly drawing on external facilitation. Organisations assess their achievements against sub-criteria, identify shortfalls, develop a change programme to work towards excellence, implement change and review its effectiveness. The underlying logic is simple: by improving how the organisation operates (the 'enablers'), excellent results will follow. The acronym associated with this logic is

RADAR:

Determine the financial, operational and stakeholder **[R]**esults aimed for; Plan a sound **[A]**pproach for achieving the proposed results; **[D]**eploy the approaches systematically to

achieve integration; **A**ssess and **R**evue the approaches by undertaking regular measurements, leading to learning and change.

The UK government has explicitly commended the European Foundation for Quality Management Excellence Model (EFQM EM) as a framework for delivering continuous improvement under the new health care quality agenda (NHSE 1999). In the context of clinical governance, supporters have championed the EFQM EM as a tool for undertaking baseline assessments (Holland and Fennell 2000), and as a framework for implementation (Jackson 2000). Supporters suggest that its suitability is due to its potential to absorb current initiatives; emphasis on self-assessment; development of leadership potential; encouragement of 'best-in-class' benchmarking; and ability to foster a quality improvement culture (Jackson 1999, Stahr 2001).

While the high face validity, simplicity and compatibility with the structure-process-outcome approach have ensured the relative popularity of the EFQM EM in health care applications it remains a generic model (Klazinga 2000). Health care providers may use the EFQM EM for voluntary self-development or seek an external European Quality Award; this latter approach could be considered to be external and formative in the Boland and Fowler matrix (Fig. 4.1.2).

4.2.2.2 Benchmarking processes

Benchmarking has been defined as the search for best practices that lead to superior performance (Camp 1989), and is essentially a technique for learning from the practices of

others by comparing current processes with those used by leading organisations in order to identify and implement improvements. The emphasis is thus on the study and transfer of exemplary practice rather than the general comparison of one organisation with another (Jones 2001; Modernisation Agency 2003). The principal advantage claimed for the approach is its ability to identify practices proven to increase service quality (Dence 1995).

4.2.2.3 The professional perspective: *Visitatie*

Visitatie consists of systematic site visits by selected peers, resulting in publication of an assessment report. Typically organised by medical specialities and especially popular with in the Netherlands, *visitatie* programmes are developed and executed by and have their roots in the profession, placing the emphasis on the knowledge, skills and attitudes required for clinical performance (Klazinga 2000). While the functioning of a group of specialists within a particular hospital is the most common approach, some programmes have developed more broadly to include organisational issues.

Visitatie principles have influenced the composition of the Commission for Health care Improvement (CHI) review teams, which consist of peers from a range of professional groups, including medicine, nursing and professionals allied to medicine, as well as managers and service users. The approach was also utilised by the former West Midlands Regional Office (WMRO) in 1999 in 'pre-CHI' developmental visits aimed to encourage health care organisations to develop their clinical governance infrastructures (Freeman 2000).

4.2.2.4 Modernisation agency

The Modernisation Agency (MA) was announced in the NHS Plan (2000) and forms part of the means of implementing the plan. The MA draws together leadership, management development and organisational change capacity to support the modernisation agenda, and its chief executive, David Fillingham, characterises the agency's remit as performance improvement (Smith 2001). It is essentially an umbrella organisation that integrates the work of a number of agencies, such as the National Patient Access Team, National Primary Care Development Team, Clinical Governance Support Unit and Leadership Centre. These agencies run multiple programmes, including waiting times and service redesign initiatives. The agency evaluates the impact of its programmes, including their impact on local service provision as well as staff and patient satisfaction.

4.3 Internal approaches

4.3.1 Internal summative approaches

While internal summative approaches imply the collection of performance indicators for internal praise or sanction, in the context of a strong external performance management system internally collected data is typically aligned with the information needs of the external framework. Thus within the current regime trusts are required to return information on indicators contained within the PAF, and while data collection is internal, the focus is on the external performance management framework.

4.3.2 Internal formative approaches

4.3.2.1 Statistical process control (SPC)

Based on Walter Shewhart's work on the economic control of variation, SPC uses a graphical method known as a control chart to explore causes of variation in processes, which may be either common or special cause. Common cause variation is intrinsic to the process, and thus in order to decrease variation the process itself needs to be changed. In contrast, special cause variation results from extrinsic factors, and its reduction requires identification of and action on the special cause. In order to differentiate between common and special cause variation, Shewhart recommends plotting mean values and control limits at 3 standard deviations above and below the mean (Mohammed *et. al.* 2001). Values falling outside the control limits are outside of the variability that one might reasonably expect from a population distribution- they are thus identified as due to special causes. Variation within the control limits is common cause i.e. within the natural variation that one might reasonably expect within a dynamic system.

The acceptance of random variation avoids the problem of lack of sensitivity and specificity on in performance rankings (Adab *et. al.* 2002), and provides a basis for action to either change underlying processes, or identify and eliminate special causes. When applied to benchmark data from multiple sites, SPC provides excellent comparative information for service improvement activity.

4.3.2.2 Specific measures of clinical governance

4.3.2.2.1 Clinical governance self-assessment and development guides (WMRO)

Produced by Partnership for Developing Quality (PDQ) and funded by the former West Midlands Regional Office (WMRO), the West Midlands Clinical Governance Self-Assessment and Development Guide (NHSE West Midlands Region 1999) provided an early attempt to measure progress in clinical governance. The guide was based on an earlier round of visits to local NHS trusts to support development of clinical governance portfolios as a precursor to CHI visits. The measure covers five areas of clinical governance: organisational arrangements; staff development; knowledge and information management; clinical effectiveness; and learning from risk management. In each area key components are identified. For example, key areas for organisational arrangements include board level leadership, strategic links for sharing good practice and arrangements for local delivery of national and regional guidance.

In each key area, indicators of good practice are identified under the headings structure, process, outcome, culture and knowledge. For example, the first key area under organisational arrangements, board level leadership, lists the importance of an integrated strategy reflecting staff views under structure, and the receipt by the board of regular reports from all areas of the organisation which influence clinical governance. The guide recommends compilation of a portfolio of supporting evidence around each of the indicators.

The principal difficulty with the approach relates to its insensitivity to variance within organisations. The focus of the assessment framework is on gathering data to assure an external agency that 'all is well', and the use of a portfolio of evidence encourages those responsible for data collation to provide the best possible gloss. There is pressure to avoid exploration of variance in achievement across different divisions / service areas / teams and the schedule itself encourages a global response.

4.3.2.2.2 CHI self-assessment schedules

Developed after the OPCG and CGCQ between September 2001 and October 2002, the CHI self assessment tools are designed to identify development areas at three distinct layers within organisations: clinical team, senior management, and corporate management. The first concentrates on aspects of patient experience and the latter two on strategic capacity. Each tool contains a series of statements to encourage reflection on how things currently operate, together with guidance on use, six-point rating scales and a section for noting key action points. The combination of space for recording discussion and summary scores for current achievement against statements and amount of influence over improvements allows staff to identify and prioritise an improvement agenda. Differentiated tools are available for acute, ambulance, mental health, combined and NHS direct, and at time of writing in summer 2003 a primary care tool is under development.

4.4 The need for further measures

Given the extent of measures and regulatory systems detailed above, is there any need for yet more measures? Or are they, in the well-known absurdist phrase, about as necessary as bicycles for fish? The problem with existing measures is that they are mainly summative, and thus pose difficulties in applications related to quality improvement. Those few that are formative are either general techniques for service quality improvement, such as benchmarking and statistical process control, or partial in their coverage of concepts associated with clinical governance, or of unknown reliability and validity. Thus there is a need for measures of progress in clinical governance with known reliability and validity that cover all of the concepts within clinical governance. Two such measures are proposed. The first is for a measure of clinical governance climate at the level of staff perception, and the second for a measure of organisational progress at division / directorate and board level.

4.4.1 The case for a measure of clinical governance climate

As explored in chapters one and two, the long-term agenda of clinical governance concerns quality improvement, and implies far-reaching cultural change. Indeed, official policy documentation explicitly describes clinical governance as being principally concerned with cultural transformation:

“Above all clinical governance is about changing organisational culture in a systematic and demonstrable way, moving away from a culture of blame to one of learning so that quality infuses all aspects of the organisation's work.” (NHSE 1999: p. 5)

It would be difficult to be against the general thrust of such a statement- who after all would choose a culture of blame over one of learning? However, the use of the term organisational culture in connection with health care begs difficult questions about its nature and the extent to which it may be transformed (Davies 2000). Despite the unproblematic use of the concept by management gurus (Peters and Waterman 1982), there are a number of conceptual difficulties including the contested nature of organisational culture in the literature, its implied homogeneity, and the limits to wholesale cultural transformation.

Contested nature of organisational culture

The organisational culture literature is conventionally divided into two strands. The first conceptualise culture as an *attribute*, something that an organisation *has*, which may consequently be identified and manipulated. In contrast, the second strand regards culture as something that an organisation *is*, a base metaphor for describing an organisation. Closely associated with postmodernism, the second strand does not presume the existence of a thing called an 'organisation' with an external boundary against its environment. Rather, the emphasis is firmly on a continuous process of organising- organisational cultures are multiple and in a constant state of flux in which meanings are negotiated and organised. Bate (1994) summarises the postmodern perspective well:

"Cultures are not physical entities. They are complex social phenomena produced interactively not biologically."
(Bate 1994: 136-7)

The distinction between culture as attribute and culture as base metaphor has important policy implications. Culture as attribute holds the possibility of the manipulation of culture in line with organisational objectives, which is highly desirable to those charged with implementing the NHS reforms. In contrast, if organisations are seen as cultural systems then change agents have fewer levers to influence the formation of desirable cultures, and need to proceed with greater caution and sensitivity.

Implied homogeneity

The model of organisational culture contained in the NHSE statement on clinical governance is integrative (Martin 2002) i.e. it implies that health care organisations' possess a unified culture (singular), which needs to be shifted from a present undesirable state to another, more desirable state. Yet, the implied homogeneity is inconsistent with characterisations of NHS organisations as consisting of multiple professional cultures. Indeed, the mix of managerial and professional cultures in which multiple values co-exist alongside power imbalances among the various stakeholders has been seen as they key distinctive feature of the NHS (Harrison *et. al.* 1992; Degeling *et. al.* 2001; Pollitt 1993; Dawson 1999). In such a context, there is a risk that attempts to develop a strong unitary culture may overlook the richness within the pre-existent culture(s) and the multiple, complex meaning-making processes (Hawking 1997).

Limits to cultural transformation

The management and clinical governance literature contains examples of claimed 'makeovers' of organisational culture (Bernick 2001; Shirley 2000). Yet, if the

assumptions of culture as attribute and cultural homogeneity are challenged, the potential for management interventions to have a transformative effect upon organisational culture(s) is reduced. Such conceptual refinements take on even greater significance in the context of a complex interdependency of individuals and teams, in which each can only reach their objectives by relying on others seeking to achieve different objectives (Iles and Sutherland 2000). It is for these reasons that Davies *et. al.* (2000) urge a cautious reassessment of the possibilities for cultural transformation in NHS organisations involving sensitivity to sub-cultural diversity, honouring current achievements, and consideration of the needs fears and motivations of diverse staff groups.

While there is still scope for leading change it is a more contingent enterprise. Effective leadership in such a context requires sensitivity to the needs of multiple stakeholders, and the ability to confront discrepancies between the current dominant culture and the demands of those at the boundaries of the organisation such as patients / clients, stakeholders or political agencies (Hawkins 1997). Failure to address the implicit political dimensions of organisational change in such a context runs the risk of the intervention becoming a crude attempt to extend the power of management at the expense of others:

“Consultants can fail to see the socio-political dimensions which can lead to a culture change programme being an attempted instrument of power and control by the dominant leadership group, which if unilateral and insensitive to the cultural complexity, will inevitably lead to a reaction and often rejection by those on the receiving end of the new wave of cultural imperialism.”
(Hawkins 1997: p. 425)

Thus crude attempts at managing culture may all too easily degenerate into the enforcement of espoused behaviours, especially where management attempts to modify culture are driven by a desire to control (Ogbonna and Harris 1998).

Notwithstanding the above difficulties, the importance of a participative, flexible and innovative organisational climate to successful implementation of quality improvement initiatives is beginning to be understood (Shortell *et. al.* 1995; Scott *et. al.* 2003a). Given the context of multiple professional staff groups, such a climate requires an emphasis on team working across professional and organisational boundaries supported by continuous learning (Degeling *et. al.* 2003). Health care managers would find it extremely useful to be able to gauge clinical governance climate, in order to identify priorities for organisational development activity and evaluate the impact of programmes designed to engender multidisciplinary work to address specific shortfalls. Indeed:

“Whilst it is important to acknowledge that some observers contest the existence, nature and impact of culture, this should not inhibit empirical research into an issue that has great potential as a lever for quality improvement and an aid for understanding the management of change in health organisations.”
(Mannion *et. al.* 2003: p.8)

Consequently, there is scope for the development and use of a quantitative measure of the lived experience of staff related to clinical governance, sensitively used within the context of organisational development programmes.

Distinctions between organisational culture and organisational climate

The literature on organisational culture and climate and the distinctions between them has been reviewed and discussed several times (James and Jones 1974; Schneider 1975; Glick 1985; Schein 1992), the most helpful summary of changes in the use of these concepts over time provided by Guldenmund (2000). Originally coined to refer to a global concept underlying organisational events, the term climate has over time come to mean the overt manifestations of culture within an organisation. The more global concept has in turn become to be referred to by the term organisational culture. Simply stated, organisational culture expresses itself through organisational climate. Guldenmund highlights the way that these conceptual changes are reflected in the way that 'climate' and 'culture' have come to be operationalised. Thus, organisational climate is typically conceived as a distinct formulation of a limited number of dimensions, surveyed by the use of self-completion quantitative questionnaires, similar to the psychometric study of attitudes. In contrast, organisational culture is determined phenomenologically, through the lived experience of a range of organisational actors, and thus explored qualitatively drawing on depth interviews and observations (Figure 4.4.1.1).

There is broad agreement within the literature that organisational cultures are multi-layered, consisting of a central core that is overlaid with more superficial and easily influenced outer layers. While the detail is disputed, there is consensus over the general ordering, from basic assumptions at the core, through cultural artefacts and symbols / codes of conduct / procedures (Hofstede 1991; Schein 1992). Schein's model, widely cited in the literature,

Figure 4.4.1.1: Conceptual distinctions between ‘culture’ and ‘climate’ research

	Culture	Climate
Theoretical underpinning	<i>Phenomenological</i>	<i>Attitudinal</i>
Object of study	<i>Beliefs and values</i>	<i>Behaviour or expression</i>
Academic discipline	<i>Anthropology</i>	<i>Psychology</i>
Data type	<i>Qualitative</i>	<i>Quantitative</i>
Data collection methods	<i>Observation / interview</i>	<i>Questionnaire</i>

specifies three layers: cultural artefacts, espoused values (attitudes), and basic assumptions. Artefacts are the outermost layer, and are the most visible manifestations of culture, such as its rituals and rewards. The second layer, espoused values, refer to those values used to justify behaviour and constitute the grounds on which alternative courses of action are justified. At the core lie assumptions, the unspoken and often unconscious beliefs and expectations shared by individuals. Organisational climate thus refers to the middle layer, espoused values. These conceptual distinctions form the basis of the contested nature of organisational culture, between those who view organisational culture as an attribute that an organisation *has* which may be manipulated (‘climate’); and those that view it as something that the organisation *is* (‘culture’), defining the whole character of organisational life.

Guldenmund (2000) summarises the characteristics, and difficulties pertaining to the exploration, of organisational climate. Firstly, they are abstract concepts rather than concrete phenomena, which provides space for disagreement over definition and

operationalisation of constructs. They also possess multiple dimensionality, comprising several composite variables, and the labelling of dimensions can be a source of disagreement between researchers. Thirdly, climates are shared by groups of individuals – they are reciprocal and integrating. Consequently, large organisations may contain multiple climates within different divisions / departments and groups, so that the issue of aggregation becomes important. Climates also consist of various aspects, so that several climates may be distinguished within an organisation such as ‘safety climate’ (Coyle *et. al.* 1991), or ‘innovative climate’ (Ekvall 1983), or indeed clinical governance climate. These distinctions are pragmatic in that they narrow the concept in order to concentrate attention on the central topic of interest. Finally, it provides a frame of reference for behaviour- ‘the way that things are done around here’.

Organisational climate thus principally relates to shared beliefs about the common practices within organisations and may be summarised as:

“... a relatively stable, multidimensional, holistic construct shared by groups of organisational members that supplies a frame of reference and which gives meaning to and/or is typically revealed in certain practices.”
(Guldenmund 2000: p. 225)

Extant measures of organisational culture / climate

In a recent systematic review of quantitative measures of organisational culture / climate that had either been validated and used in health care settings or appeared to have potential for use in health care settings, Scott *et. al.* (2001; 2003c) identified a total of 13

instruments (Table 4.4.1.2). These measures varied in the extent to which they were conceptually grounded or pragmatically derived; global in scope or limited to specific sub-components of organisational culture; or the extent of validity and reliability resting. The authors helpfully distinguish between typological and dimensional approaches; the former referring to assessments resulting in assignation of one or more ‘types’ of culture, and the latter describing culture in terms of position on a number of continuous variables.

Table 4.4.1.2: Quantitative organisational culture measures After Scott *et. al.* (2001; 2003)

Typological approaches	Competing Values Questionnaire Harrison’s Organisational Ideology Questionnaire Quality Improvement Implementation Survey
Dimensional approaches	Organisational Culture Inventory (OCI) Hospital Culture Questionnaire Nursing Unit Culture Assessment Tool Practice Culture Questionnaire MacKenzie’s Culture Questionnaire Survey of Organisational Culture Corporate Culture Questionnaire Core Employee Opinion Questionnaire Hofstede’s Organisational Culture Questionnaire Organisational Culture Survey

The authors argue for the pragmatic selection of instrument based on the purpose and context of assessment. Thus, a global view requires a comprehensive multi-dimensional instrument such as the Organisational Culture Inventory (OCI), while focused studies may require more focused tools, such as the Nursing Unit Cultural Assessment Tool to examine the culture of a specific professional group.

While the review identified a number of global measures, it did not identify any measure of the sub-components of clinical governance climate. Existing measures were either global measures of organisational performance, or else related to specific professional / organisational sub-groups. Consequently, there is space for a dimensional measure of clinical governance climate, relating to the lived experience of staff. This thesis proposes the development and use of such a measure, the Clinical Governance Climate Questionnaire (CGCQ).

Methodological difficulties associated with aggregation

The major methodological problem associated with quantitative measures of climate concerns the appropriate level for score aggregation. The problem refers to the extent to which individual perceptions can be combined to provide information about a whole organisation or constituent organisational units such as departments or directorates. This is an important methodological issue, as seemingly obvious aggregation levels may carry the potential for erroneous inference. Typically, recourse is made to statistical tests for levels of homogeneity in order to justify aggregation, so that analysis of variance designs are used to ensure that the differences between aggregated groups are greater than the differences within them (Glick 1985).

Most empirical explorations of organisational culture / climate work from the premise of homogeneity- that culture is nearly universally shared within organisations. Those few empirical studies that directly address the issue found that perceptions were mediated by nationality, demographic characteristics and position in hierarchy (Hofstede *et. al.* 1990;

Keeton and Mengistu 1992; Helms and Stern 2001). Given the widely noted professional tribalism in health care organisations (Degeling et. al. 2003), perceptions of clinical governance climate are likely to vary by professional sub-group, and for this reason sampling was designed to allow subgroup analysis. Consequently, the study makes a distinct methodological contribution concerning the mediation of perceptions of clinical governance climate within organisations by professional group affiliation. Do perceptions vary between organisations regardless of professional affiliation, so that there is a distinctive ‘trust climate’ for clinical governance? Or do perceptions vary between professional groups regardless of organisation, so that there is a distinctive ‘professional climate’ for clinical governance? These methodological issues are addressed within the results and discussion sections of chapter six.

4.4.2 The case for a measure of organisational competency in clinical governance

Health care trusts charged with implementing clinical governance require indicators of progress in order to identify areas of weakness and trace the impact of their implementation strategies over time. The CGCQ above provides detail of the lived experience of staff at the ‘shop-floor’ level of an organisation and its congruence with good clinical governance. However, given their responsibilities for implementing policy, the perceptions of board- and directorate-level staff on progress in clinical governance are also valuable. A number of such corporate level measures of progress in clinical governance are available, such as that developed by the former Regional Offices in the West Midlands described above (refs. above). However, the proposed organisational progress in clinical governance (OPCG) schedule differs from previously available measures in three important respects. Firstly, the

emphasis is on internal development rather than external assurance. Secondly, it emphasises a variance approach to data collection, rather than a portfolio of evidence. Finally, the proposed measure makes use of a competency framework consisting of multiple statements of achievement, rather than seeking opinion on achievement against abstract concepts.

(a) External assurance Vs Internal development

External assurance approaches assess the functioning of an organisation against explicit criteria. They require the organisation to supply evidence of its compliance with a series of attainment targets. Given the need to perform well against others, there are implicit management incentives to present information in the best possible light. Indeed, the unintended consequences of performance indicator systems were explored in detail in chapter three above. In contrast, the proposed organisational measure is designed for internal use by health care organisations. The emphasis is on generating insights for future development, rather than assuring an external agency that 'everything is fine around here'.

(b) Portfolio Vs variance data collection

Portfolio and variance approaches to data collection each involve the scoring of an organisation against specified levels of attainment. Portfolio methods, such as that designed by the former WMRO, require a single score in each area of interest, validated by a portfolio of supporting evidence (see 4.4.2.1). In contrast, the variance approach utilised in the OPGC schedule dispenses with a portfolio of supportive evidence, and instead of collecting a single score for each area of interest, scores are collected from multiple

respondents. Samples of respondents are drawn from two distinct populations, reflecting two tiers within the organisation: board directors, and clinical team managers. Board directors and clinical team managers are asked to identify the level at which they believe the organisation is operating on a series of clinical governance related tasks. Consequently, data reflects a range of opinion about progress in each task. Measures are thus indicators of perceived progress, and may be used to stimulate further discussion about appropriate actions for development.

Box 4.4.2.1: Data collection approaches

Portfolio

Typically used in assurance approaches, portfolio data collection requires a single rating of a level of attainment against specific standards. Supporting documentary evidence is requested. The difficulty is that there is little provision for examining differences in perceived attainment across the organisation. There is also pressure to 'game' the indicators and present the organisation in the best possible light.

Variance

Multiple respondents identify the level of attainment of their organisation or team. Differences of opinion are visible, and are used as the basis of discussion on areas of strength and weakness. This is used to influence future planning and to promote learning across the organisation.

c) *Competency framework with active tasks*

In order to discriminate between high and low performance in clinical governance operation, the OPCG applies a 'competency' framework to a series of tasks related to the operation of clinical governance. The competency framework reflects an 'integrated' conception of competency (Hager and Gonczi 1996). This model overcomes the

difficulties involved in viewing competence as either 'task completion' or 'desirable attribute'. The former tends to omit the broader aspects of competent performance such as planning and reaction to contingencies, due to the need to specify specific tasks. The latter suffers from the problem that attributes assessed outside of a specific work context bear little relation to future performance; skills application is important but remains unmeasured. The integrated model examines the knowledge and skills displayed in the context of a series of tasks. The measure consists of tasks that need to be performed well in order for clinical governance to deliver its promise, aggregated into conceptual areas related to clinical governance. For each task, items that specify functioning at different levels of competency are detailed.

It is particularly important that the competency framework is based around task completion, as it anchors the framework in the world of action rather than abstract concepts.

Consequently, the measure does not require respondents to rate the success of the organisation in achieving abstract concepts such as knowledge management. Rather, the concept is itself rooted in successful practices such as clinical change being based on assessments of evidence, or the consistent use of information to inform the quality improvement agenda. This is in marked contrast to those frameworks that request that managers provide evidence of knowledge management, or organisational learning, within an organisation (NHSE West Midlands region 1999).

Conclusions

At national and local level there are many options for assessing progress in clinical governance including performance management, quality management frameworks and accreditation. Negative, external performance measures associated with clinical governance include performance ratings ('stars'), Performance Assessment Framework (PAF) indicators and audit tools associated with National Service Frameworks (NSFs). Such assurance-related systems suffer from the conceptual difficulties and risk the perverse incentives detailed in chapter three. While there are a number of internal and formative approaches available, they tend to be either generic, such as benchmarking and statistical process control, or measures with unknown validity and reliability or an underdeveloped conceptual base. While many generic measures of organisational culture / climate could be used to assess progress in clinical governance, they tend to be related to broader issues of organisational performance and do not cover all of the areas related to clinical governance. Consequently, there is scope for specific measures of progress in clinical governance.

As perceptions of reality are material and constraining to social actors, impressions of what things are like are important, therefore a climate measure would provide valuable information on the perceived reality of staff. There are attendant difficulties, most notably the existence of distinct professional perspectives on aspects of health care quality reform and difficulties over appropriate levels of aggregation. However, the extent of convergence of perceptions of how things operate could be seen as a marker for a corporate approach to clinical governance.

CHAPTER 5
MEASURES OF CLINICAL GOVERNANCE I:
The Clinical Governance Climate Questionnaire (CGCQ)

Introduction

The CGCQ was designed to measure individual staff experience of the climate of clinical governance within health care organisations. This chapter begins with a précis of the research used in its development, before turning to a more detailed exposition of the methods used at each sequential stage in the programme. Methods included a brief literature review to identify the boundaries of indicator development, followed by a series of focus groups to qualitatively develop indicators associated with staff perceptions of nine domains of clinical governance climate identified in the prior review. The content validity of these draft indicators was explored via a questionnaire study, which was then followed by the administration of the indicators to a development sample for item reduction, assessment of internal consistency and inductive generation of measurement domains via exploratory principal factor analysis. Finally, the construct and discriminant validity of the empirically derived factors were assessed by comparing scale scores on the new factors with scores on pre-existent measures of organisational culture and burnout, and assessing the congruence of the observed results with expected hypothesised patterns.

An overview of the normative focus group literature revealed tensions principally related to sampling, group composition and the transferability of results, which are a consequence of the divergent research traditions of various commentators. Injunctions are partly heuristic

devices, and much of this literature overtly encourages transgressive readings, challenging researchers to make appropriate choices in the light of specific research questions and settings. Following an examination of the content validity of the draft item indicators, the normative literature on the application and advantages of factor analytical approaches is reviewed in the context of the present study. The use of an exploratory principal factor model is defended on grounds of the need explore the underlying factor structure of clinical governance climate and build theory with application beyond the current dataset.

5.1 An overview of CGCQ development methods

The CGCQ was developed through a programme of four linked research projects, to ensure that the resultant scale was relevant, couched in the language of health care staff and with known reliability and validity. Detailed below, these four phases are summarised in Table

5.1.1

(a) Literature review to identify potential climate measurement domains

A brief review of *The new NHS: modern, dependable* (1997), discussion document *A First Class Service* (1998) and guidance document *Clinical Governance: quality in the new NHS* (1999) identified nine potential domains of the lived experience of clinical governance climate, which formed the boundaries for subsequent development work. Summary information on identified domains, rationales for inclusion and potential sub-domains are identified in section 5.2 below. While helpful in framing the development of measurement

Table 5.1.1: CGCQ development research programme

Stage	Aim	Process
(a) Literature review	To identify potential climate measurement domains	1. Collate primary clinical governance literature 2. Review identified literature 3. Inductively generate potential climate measurement domains
(b) Qualitative exploration of potential measurement domains	To explore the attitudes of health professionals towards the identified domains	1. Identify participants 2. Conduct focus groups and discuss each domain 3. Content analysis on the range of expressed opinion 4. Develop an item pool of statements based on the content analysis
(c) Content validation of item pool	To review the internal validity of the item pool empirically with an expert sample	1. Identify a notional expert sample 2. Postal survey to all operational clinical governance managers in the West Midland region, to identify problems with the draft items 3. Generate revised item pool
(d) Item pool reduction via development sample	To reduce the item pool; assess the reliability (internal consistency) and external validity (discriminant and construct) of the items; to describe an underlying factor structure	1. Postal questionnaire to a development sample in 5 NHS trusts (n= approx. 200 in each trust) 2. Factor analysis of responses to assess the underlying structure of the items 3. Discriminant and construct validity testing of the scales to ensure they behave in anticipated ways

indices, the domains identified in the review were conceptually indistinct and confounded: e.g. 'responding to incidents' and 'managing risks' potentially share many elements of good practice such as open reporting, record keeping and action on perceived problems. Given the lack of conceptual clarity and potentially unbounded nature of further literature reviews in each of the nine domains, the decision was taken to inductively build a conceptual framework for clinical governance climate. This required the development of indicators of the domains identified in the review and an empirical exploration of relationships between these indicators in order to derive common latent factors- such common latent factors constitute an empirically derived conceptual framework for clinical governance climate.

(b) Qualitative exploration of potential measurement domains to inform indicator development

The domains identified above were explored qualitatively in a series of focus groups with NHS staff. Nine focus groups were conducted, in four separate health care trusts, with between six and nine members in each. Participants included consultant and junior medical staff, nurses, PAMs and senior managers. The focus group topic guide required participants to identify how each domain would be addressed in an ideal world, and how current arrangements in their trust differed from that ideal. The data thus contained positive statements about the domains from the ideal scenario, and negative statements from the sections dealing with the shortfall from the ideal in current practice. This method of collection facilitated the inductive development of an item pool of statements with positive and negative wording within each identified domain, based on a simple content

analysis of the qualitative material, and reflecting the lived experience of health care professionals.

(c) Content validation of the item pool

The adequacy of the draft item pool was empirically tested by means of a postal content validation study. In this procedure, notional experts reviewed the attitudinal statements in the item pool and rated their utility in describing the domains. Respondents were encouraged to suggest new indicators where they feel that the initial items were inadequate, and also to suggest alternate wordings for items that they believed potentially confusing. In this instance, the notional expert group chosen consisted of all health care trust managers with lead operational responsibility for clinical governance in the West Midland region. Respondents were generally positive about the item pool, and were able to suggest a number of improvements to individual items.

(d) Item pool reduction and external validation

On completion of the content validation exercise, the final development round required administration of the draft measurement scales to a large development sample of approximately 1000 staff from four NHS trusts within the West Midlands region. The purpose of the final development sample was to reduce the number of items in the pool, describe an underlying factor structure, and assess the reliability and validity of the retained items. These aims were addressed statistically by a combination of factor analysis and construct validation exercises. Factor analysis enables the item pool to be reduced, tests the internal consistency of the items and empirically describes a factor structure for the item

set. Construct validation assesses the ability of the measure to conform to hypothesised assumptions about its behaviour against related constructs, and discriminant validity assesses its ability to show anticipated differences between groups. In the present study, construct validity was assessed against the Maslach Burnout Inventory (MBI) and Competing Values Questionnaire (CVQ). The former measures the extent of 'burnout' and the latter the extent to which an organisation is hierarchical, rational, developmental or 'clan' based. Both MBI and CVQ scores were expected to be associated with clinical governance climate. In the case of the former, one would expect a positive clinical governance climate to be less likely to develop burnout among staff. In the case of the latter, one would expect that organisations scoring highly on developmental and 'clan' based domains would be more likely to develop positive clinical governance climates than those scoring high on hierarchy.

5.2 Literature review to identify potential climate measurement domains

The White Paper *The new NHS: modern, dependable* (1997), discussion document *A First Class Service* (1998) and guidance document *Clinical Governance: quality in the new NHS* (1999) were reviewed in order to set the boundaries for subsequent development work and identify relevant potential measurement domains. In this application, the emphasis was firmly on the lived experience of staff in areas related to clinical governance and consequently there was little point in selecting domains related to clinical governance structures. Such a domain is more suited to the measure of organisational progress (OPCG) and is covered in more detail in chapter 7 below. As a result of the review, nine potential domains related to the lived experience of clinical governance climate were identified (table

5.2.1). Domains concern the operation of multi-disciplinary teams, development of continuous quality improvement, appraisal and staff development, and the active management of clinical / non-clinical risks and adverse events.

The first, multi-disciplinary team operation is particularly important given its status as the primary mechanism of service delivery; clinical governance is mediated through team performance. Unless teams are able to work together effectively, services will be of poor quality and there is little prospect of service improvement. Related literatures include team functioning (Guzzo & Dickson 1996; Kivimaki *et. al.* 1999), innovation (Amabile 1988), and quality improvement (Firth-Cozens 1998). Secondly, continuous quality improvement (CQI), notions of a learning organisation, organisational development and use of evidence based practice are important given their potential to identify opportunities for improving service quality, managing change processes, and sharing improvements across organisations. Related literatures include structured problem-solving approaches and the use of cross-functional teams in problem solving (Shortell 1995; Gerowitz 1998), learning how staff groups in organisations learn (Davies 2000; DoH 2000), for vision, leadership and motivation in quality improvement (Garside 1998), and skills in searching, appraising and applying research (Newman 2000). Staff development and appraisal opportunities are identified in the clinical governance documents as providing opportunities for continuous development of professional skills in line with organisational requirements, as well as ensuring that professionals are fit to practice and can report problems without fear of punishment (DoH 2001). The experience of annual review and development opportunities

Table 5.2.1: Potential measurement domains for clinical governance climate

Domain	Rationale	Potential sub-categories	Key ref(s)
Multi - disciplinary team operation	MDTs are the mechanism through which services are delivered. Clinical governance is mediated through team goals and performance. They are the arena in which change needs to be effected, so are critical to making clinical governance 'real'	Vision; participation; task orientation; innovation support Common goals; clear objectives; interpersonal skills; mutual accountability Clear roles for individuals.; assessment and feedback [individual and team]; reflexivity and adaptation to new context	Amabile 1988 Ekvall et al 1999 Firth-Cozens 1998 Guzzo & Dickson 1996 Isaksen et al 1999 Katzenbach 1993 Kivimaki et al 1997 Kivimaki et al 1999
Quality improvement processes in place and working	CQI is a useful way of identifying opportunities for change and managing the change process	Focus on underlying processes as causes of failure; use of structured problem solving approaches; use of cross-functional teams to address problems; employee empowerment to identify problems / solutions; explicit focus on internal / external customers	Shortell 1995 Gerowitz 1998
Learning organisation	Opportunities for learning and development identified and shared organisation-wide	Single- , double- and meta-loop learning	Davies 2000 DoH 2000
Approaches to organisational change	The way that the organisational change is pursued affects the potential for real change	Vision; leadership; motivation; investment in quality improvement;	Garside 1998
EBP skills and implementation	To ensure continual updates to practice, teams need to have skills in searching, appraisal and implementation of evidence / protocols /	Recognising the need for information; conducting / instructing searches; appraisal of papers; application of results ¹	Newman 2000
CPD	Identification of development in line with needs of the organisation, and provision of appropriate opportunities, are essential for continual improvement	Regular identification of areas for update; linkage of professional and organizational needs; opportunities to use skills acquired in training	DoH 2001
Appraisal	Mechanisms for assessment of professional competence are required in order to assure care quality	Annual review; use of agreed performance indicators to identify problem areas; whistle-blowing arrangements [avoiding blame]; job description changes	DoH 2001
Responding to incidents	Clinicians need to be able to record and examine incidents in order to learn for the future and avoid recurrence.	Openness and encouragement of reporting; lack of blame; use of a broad approach in investigations [team, indiv., tasks, work environment]; changes to systems where necessary	Vincent 2000 CHI 2000
Managing risks	A key item in the legislative programme of CG	Existence and use of risk management systems; integration with audit and R&D programme	DoH 2000 Vincent 2001 Walshe 1998

is thus highly important, as it provides an opportunity to ensure that resources are made available to support assurance and improvement initiatives. Risk management and responding to adverse incidents are important items in the legislative programme of clinical governance, involving recording and analysing patterns, integrating risk management, audit and research programmes, encouraging openness, a systems approach to investigation and service redesign (Vincent 2000; Walshe 1998; DoH 2000; CHI 2000).

These domains are helpful to the extent that they identify areas of practice that are considered important in clinical governance and can frame the development of measurement indices. However, they lack conceptual clarity and potentially share many elements of good practice between themselves. For example, 'responding to incidents' and 'managing risks' potentially share many elements of good practice, such as open reporting, record keeping and action on perceived problems. Similarly, quality improvement and organisational development domains share the need for openness and ability to make service improvements. Given the lack of conceptual clarity and potentially unbounded nature of further literature reviews within the domains, the decision was taken to build a conceptual framework inductively. This involved a qualitative exploration of the identified domains to identify indicators of a climate conducive to clinical governance, followed by inductive, empirical theory building using exploratory factor analysis.

5.3 Qualitative development of the CGCQ item pools

This section outlines a qualitative exploration of the attitudes of NHS staff towards nine domains related to clinical governance, identified from a prior review of the clinical

governance literature. Focus groups were used to explore the domains from staff perspectives, and generate draft indicators of clinical governance climate.

5.3.1 Focus groups

Focus groups are a particular form of group interview, exploiting group dynamics to generate rich qualitative data. They are an ideal method for generating statements for indicator item pools. While there is broad agreement over their form, a close reading of the normative literature reveals differences in prescription, specifically related to sampling, group composition and the transferability of results. Such tensions are a result of the divergent research traditions of commentators. Differences in prescription should not be overstated however, as they are partly heuristic devices. Indeed, normative prescriptions are typically contingent, encouraging researchers to engage with theoretical and pragmatic issues in focus group study design. Focus groups are a useful method for gathering opinions. Notwithstanding normative injunctions against heterogeneous groups, they are able to provide information useful for scale development. Researchers need to balance methodological rigour with the practicalities of conducting research within complex host organisations.

Focus groups: definitions

There is broad agreement within the normative literature on the basic form and function of focus groups (Table 5.3.1.1). Common attributes include the organised and focused nature of the group discussion (Powell and Single 1996), and the importance of interaction between participants (Kitzinger 1995). Focus groups are thus best characterised as a form

of group interview that places particular importance on interaction between participants. They comprise group discussion among carefully selected individuals, guided by a moderator using a carefully designed topic guide. The composition of the group, structure of the guide and location flow from a well-defined research objective; it begins with and is guided by a well-articulated purpose. They are not a substitute for problem formulation or clear thinking about a topic; rather they facilitate discussion on a topic of interest through the application of clearly formulated questions (Stewart & Shamdasani 1990).

Table 5.3.1.1: Focus group definitions

Bowling (2002: p. 394)	"Focus groups are unstructured interviews with small groups of people who interact with each other and the group leader. They have the advantage of making use of group dynamics to stimulate discussion, gain insights and generate ideas in order to pursue a topic in greater depth."
Kitzinger (1994: p. 103)	"... group discussions organised to explore a specific set of issues ... The group is focussed in the sense that it involves some kind of collective activity... Crucially, focus groups are distinguished from the broader category of group interview by the explicit use of the group interaction as research data."
Kitzinger (1995: p. 299)	"... a form of group interview that capitalises on communication between research participants in order to generate data ... focus groups explicitly use group interaction as part of the method."
Krueger (1994: p. 6)	"In summary, a focus group is a carefully planned discussion designed to obtain perceptions on a defined area of interest in a permissive, non-threatening environment ... Group members influence each other by responding to ideas and comments in the discussion ..."
Krueger (1988: p. 47)	"Focus groups have a distinctive cluster of characteristics: (1) focus groups involve homogenous people in a social interaction (2) the purpose ... is to collect qualitative data from a focussed discussion; and (3) focus groups are a qualitative approach to gathering information that is both inductive and naturalistic "
Powell & Single (1996: p. 499)	"... a group of individuals selected and assembled by researchers to discuss and comment on, from personal experience, the topic that is the subject of the research."
Stewart & Shamdasani (1990:p. 140)	"Focus groups provide a rich and detailed set of data about perceptions, thoughts, feelings and impressions of group members in their own words."

Focus groups aim to promote self-disclosure among participants, by explicitly capitalising on group dynamics in discussions. Participants are encouraged to question each other's responses, elicit clarification, and explore caveats to their statements. The method seeks to promote a safe environment for self-disclosure through careful participant selection, sensitive questioning by a moderator and the prior establishment of clear ground rules for participation (Krueger 1994).

Typically, focus groups consist of between 6-12 members drawn from a study population of interest, and sessions generally last between one and two hours until the topic has been covered to the satisfaction of participants (Stewart & Shamdasani 1990). This number of participants is small enough for everyone to contribute, yet large enough to share diverse opinions across the whole group rather than fragmenting into smaller parallel discussions (Krueger 1994). The task of the moderator is to facilitate and focus discussion on the topic of interest, ensuring that participants are able to contribute fully to the developing discussion.

Specific strengths of focus groups

The main purpose of focus group research is to draw upon respondents' beliefs, attitudes and feelings by exploiting group processes. There are many advantages to interaction between participants and, indeed, many see interaction as the key to the method (Kitzinger 1994). The idea is that group processes can help people to explore and clarify their views

and attitudes more efficiently than they could in individual interviews, and encourages participation from those who feel that they have little to say (Kitzinger 1995). The interpersonal communication between participants helps to clarify the cultural values or group norms shared by the study population. This is due to the social context of interaction, and is a major advantage of focus groups over individual interviews. In the former, analysis can draw directly on discussion between participants, whereas in the latter analysis depends much more on insights drawn by the researcher. The method additionally benefits from the increased efficiency of multiple interviewees involved at a time.

While they offer a rapid and cost-effective means of interviewing multiple respondents, the choice between focus groups and individual interviews depends on the importance of individual views in analysis. Focus groups are inappropriate where a detailed exploration of individual perspectives is required (Stewart & Shamdasani 1990; Webb & Kevern 2001), in which case individual interviews are the better option.

Choice of focus groups in the present study

Focus groups facilitate the collection of qualitative data, lending themselves to particular types of qualitative research question. The present study required identification of potential indicators of progress in clinical governance, based on the expressed opinions of NHS trust staff, and focus groups were exceptionally well suited to this task. They are useful for exploring attitudes and experiences (Morgan 1988), particularly how they develop and operate within a given cultural context (Kitzinger 1995). They have been posited as an ideal method for both generating statements for indicator item pools (Morgan 1988), and

testing the phrasing of questions for questionnaires (O'Brien 1993; Wolff *et al* 1993). Indeed, Morgan (1988) is critical of the practice of generating item pools in 'armchair' fashion rather than through contact with potential survey participants, and explicitly recommends focus groups for this purpose.

Focus groups help to capture as complete a picture of participants thinking as possible, by exploring how respondents typically talk about the topic of interest in their own vocabulary. For all of these reasons, focus groups were an ideal method for the present study. While individual interviews could have been used, there was no need to explore individual perspectives in detail, which meant that the increased efficiency of focus groups and their opportunities for interpersonal communication between participants could be exploited.

Tensions between normative prescriptions

While there is agreement on the basic form and function of focus groups, a close reading of the normative literature reveals a series of tensions between different prescriptions. These differences principally concern sampling techniques; the composition of groups; the role of interaction; and the degree of inference that may be drawn from results. The confusion is due to the adoption of the method by researchers from different traditions within the social sciences. Consequently, normative prescriptions of the method are informed by broader methodological concerns related to the logic of inquiry.

This becomes clearer if we examine a number of tensions between the prescriptions of two major contributors. Writing from a broadly subjectivist perspective, Kitzinger is primarily concerned with the situated nature of interaction between participants in focus groups. This leads her to take up distinctive positions on the value of pre-existing groups, and an emphasis on the transferability of results at the level of theory (concepts). In contrast, Krueger writes from an applied programme evaluation perspective, in which procedures are required to reduce selection bias, and increase the transferability of results from samples to broader populations. Four major tensions are discernible (Table 5.3.1.2.):

Table 5.3.1.2: Methodological tensions within normative focus group literature

	Subjectivism (Kitzinger)	Programme evaluation (Krueger)
Group membership	Pre-existing groups may be useful in providing 'naturalistic' exchanges.	Pre-existing groups should be avoided given their potential for bias. Random sampling of participants is recommended.
Homogeneity	Weak: May be helpful when participants have marked status differences. Homogenous groups may themselves lead to conformity effects and inhibit discussion.	Strong: Important that groups are homogenous with regard to important variables for sub-group comparisons.
Interaction	Strong: interaction as the central analytical resource.	Weak: interaction as useful for generating discussion on the topic of interest.
Generalisability of results	Weak: theoretical insights potentially transferable, decided by the reader.	Strong: given a number of homogenous groups, results may hold for the population from which groups are drawn, given random sampling of participants. It is up to the researcher to make a case for transferability, and to the reader to decide whether the case is made.

(a) Group membership- from random sampling to pre-existing groups

Krueger is critical of the use of convenience samples and pre-existing groups, and urges caution when using groups of people who know and work closely with each other. The first and weaker injunction is that such groups will have their own pre-existent dynamics. This poses particular problems for analysis given the influence of current relationships on contributions, the inhibition of negative observations and influence of formal and informal hierarchies (Krueger 1994).

The second injunction concerns the threat to external validity posed by convenience sampling. Krueger is interested in developing analyses with explanatory power beyond the sample, and deploys the logic of population sampling in order to avoid bias and increase the external validity of results. In contrast, Kitzinger (1994) defends the use of pre-existing groups in focus group research in order to explore how people talk about a topic. Rather than follow Krueger's logic of segmenting populations by analytically informative variables, in order to assess the external validity of emergent concepts, she explores socially situated discourses. She emphasises the situated nature of human interaction, valuing the fragments of naturally occurring data that become available for analysis. She qualifies her use of the phrase 'natural' to describe such data, accepting that it is a response to specific situations in context. For Kitzinger, the underlying logic is of reflexivity i.e. an acceptance and heightened awareness of the influences of the subjectivities of analyst and participants. This is in marked contrast to Krueger's approach, in which subjectivities are categorised as confounding bias, to be controlled and abstracted away.

(b) Homogeneity and heterogeneity in group composition

Krueger (1994) makes two injunctions against heterogeneity in composition. The first and stronger injunction is analytical, and again concerns external validity. Where the study purpose is to compare the opinions of specific sub-groups within a study population, he advises segmentation of participants into homogenous sub-groups, followed by a series of focus groups with each separate sub-group. Participants in any one focus group should all belong to the same sub-group segment. This facilitates an analysis of differences between the sub-group segments, and increases the external validity of comparisons made between sub-groups. The logic of inquiry here is broadly positivist, concerned with the external validity of inferences drawn from sample to broader population. This is not a concern for Kitzinger, as she is concerned with situated discourses rather than shared perceptions of population samples.

The second, weaker injunction refers to the need to ensure that participants are able to raise issues for discussion. Failure to segment means that important information may not be exposed (Krueger 1993). Morgan & Krueger (1993) are critical of those who compose groups that make some participants unwilling to express themselves, as this defeats the purpose of a focus group. The concern here is that too much heterogeneity will inhibit discussion, especially when there are status distinctions between participants. Kitzinger concedes that that status difference between participants may inhibit participation, but she is concerned that homogenous groups may themselves inhibit discussion due to conformity effects (Kitzinger 1994).

Despite making homogeneity normative, Krueger acknowledges a creative tension between homogeneity and heterogeneity in any focus group composition, between encouraging the sharing of information, and ensuring discussion of diverse opinions (Krueger 1988:47). He further acknowledges the difficulties involved in bringing together groups of participants, especially in an organisational context (Krueger 1993), and appreciates the need for compromise in participant selection. While driven primarily by the study purpose, he does acknowledge those practical concerns that influence participant selection.

(c) Interaction

Both identify the importance of interaction in focus groups, but for somewhat different reasons. For Krueger, interaction between participants is a helpful device for encouraging discussion on the topic; it performs a useful instrumental function in gathering data.

Kitzinger identifies a more central role for interaction, as *the* central analytical resource. It is prized for its own sake, rather than simply as an efficient way of gathering data.

Consequently, she emphasises the importance of concentrating on interaction between participants in analysis, to the point of making it a defining feature of the technique (Kitzinger 1994).

(d) Generalisability of results- degrees of transferability

Krueger makes the case for the robust external validity of well designed focus group studies. Given an adequate number of homogenous groups with randomly selected participants, study results may be transferable to the population(s) from which the groups were drawn (Krueger 1994). In contrast, Kitzinger is more wary of claims of the external

validity of findings. She asserts a weaker criterion of transferability, based on theoretical saturation of data segments. The important thing is that the researcher has refined the emergent theoretical categories in sufficient detail for the reader to make a judgement on their credibility (Kitzinger 1995).

Methodological pragmatism in the light of divergent prescriptions

Such diversity of use from within a wide range of disciplines poses a challenge for any normative 'rules' for conducting focus groups. Differences in prescription should not be overstated however, as they are partly heuristic devices. Indeed, the complexity of focus group studies is recognised by many commentators, who remain aware of the practical difficulties involved in designing and running focus groups even while arguing their case. Application of the method requires sensitivity, and considered decision-making in design and analysis. Normative prescriptions are typically contingent; encouraging researchers to engage with theoretical and pragmatic issues in focus group study designs. Researchers are often explicitly encouraged to learn from the experiences of others without being bound by them (Morgan & Krueger 1993). In the final analysis, there is a widespread commitment to methodological pragmatism. Indeed, transgressive readings of normative prescriptions are actively encouraged:

... we have tried to provide a basic guide to some of the key questions around developing focus group research. We have resisted providing definitive answers in favour of suggesting ways of considering alternatives. If our answers often seemed merely another way of saying 'well, it all depends', that is because it does! Every decision in the course of designing, conducting and analysing focus group research is interdependent. A 'one size fits all' formula would be no substitute for serious critical engagement with the political, theoretical and practical issues around group work." (Barbour & Kitzinger 1999: pp. 19-20)

The discussion above outlines the importance of a consideration of desired results in the design of focus group research, and the need for carefully defined research questions. Despite injunctions against heterogeneous groups, a range of opinion exists on their potential usefulness, and there is recognition that applied settings call for specific compromises (Morgan & Krueger 1993). Pragmatic considerations, reflecting the nature of the setting, availability of participants and resource constraints, will have an impact on any given study. The important thing is to ensure a defensible strategy based on an accommodation of a robust design on one hand, and the unique constraints posed by the research setting on the other. Decisions taken in the present study are discussed fully below.

5.3.2 Methodology

Introduction

Sixty-five doctors, nurses, paramedics and managers working at four health care trusts in the West Midland region were involved in one of nine focus groups. Participants discussed their opinion of ideal practice and current shortfalls from this ideal in nine domains related to clinical governance, identified in earlier literature reviews. A simple content analysis was used to organise data segments within each domain. Sub-themes were identified within each domain, using codes derived inductively from the data. Results are presented for each sub-theme of each domain, consisting of sample data segments to show the range of statements and the total number of data segments in each sub-theme.

Methods

As identified above, the composition of the groups and structure of the topic guide flow from closely defined research questions, filtered through pragmatic considerations concerning the setting, availability of participants and resource constraints.

(a) The research questions

The purpose of the study within the broader thesis design was to identify potential indicators of progress in clinical governance, based on the expressed attitudes of NHS trust staff. Two detailed research questions arose directly from this purpose:

1. What do NHS staff believe to be 'ideal practice' in domains directly related to clinical governance identified in an earlier literature review?
2. What differences are there between the stated 'ideal practice' and current practice in each domain?

These questions directly capture positive and negative statements about the identified domains. They thus facilitate the development of positive and negative statements relating to the work environment, in participants' own words, in each domain of interest.

(b) The topic guide

The topic guide sets the agenda for the focus group discussion, and follows directly from the research questions. Its purpose is to provide direction for the group discussion and is

thus relatively unstructured in order to avoid suggesting 'preferred answers'. It acts as a prompt to the moderator, to ensure discussion of all areas of interest.

Table 5.3.2.1: Focus group topic guide

(1) In an ideal world, how would clinical teams respond to adverse clinical events? (2) Ideally, how would clinical teams approach quality improvement? (3) Ideally, what arrangements would be made to manage clinical risk? (4) In an ideal world, what opportunities would exist to learn from practice and share insights with the rest of the organisation? (5) What foundations are required for encouraging change in the organisation? (6) Ideally, how would staff appraisal systems operate, and how would staff benefit? (7) In an ideal world, what skills and resources would teams have to develop Evidence Based Practice? (8) In an ideal world, what would multi-disciplinary working look like? (9) In an ideal world, what arrangements would be made for CPD? For each question area: What differences are there between this ideal picture and the way that things happen in your trust?

The topic guide (Table 5.3.2.1) was based nine potential clinical governance climate domains taken from government guidance on clinical governance. Domains included the operation of multi-disciplinary teams, quality improvement processes, clinical risk management, and organisational change.

(c) The number and composition of groups

As previously described, the normative literature contains a range of prescriptions of good practice in group composition, ranging from the segmentation of sub-populations (Krueger

1993, Knodel 1993) to the use of pre-existing groups (Kitzinger 1995). Pragmatic considerations apply, including the degree of homogeneity of participants in terms of background and perspective, the goal being to run only as many groups necessary to answer the research question adequately. In this study, the research questions related to 'ideal practice' and 'current practice' in domains related to clinical governance experienced by trust staff. There was no intention to make direct comparisons between different sub-groups of staff, e.g. doctors and nurses, nurses and PAMs, doctors and managers. Despite the fact that segmentation was not required for analytical purposes, the original intention was to aim for homogeneity of profession within focus groups, based on the assumption that status distinctions between members could inhibit participation. For this reason, the original study design specified nine groups; three groups each of doctors, nurses and PAMS, and managers.

(d) Recruiting participants

Notwithstanding normative injunctions against convenience samples given their potential for hidden bias (Krueger 1993), participants were selected opportunistically. Given the difficulty of gaining access to professional groups from outside of an organisation, the study involved participants drawn from a limited number of trusts previously involved in clinical governance research within the West Midlands region. This allowed relatively easy access to the required professional groups, given the goodwill developed with the sites over the course of the earlier work (Walshe, Freeman and Latham *et. al.* 2000). The recruitment process involved initial discussions with the trust Nursing and Medical

directors, to outline the study and seek support. Once this had been secured, the directors charged a senior manager with the task of acting as a contact inside the organisation.

In negotiations with contacts at the trusts, it quickly became apparent that recruitment to professionally homogeneous groups posed serious problems across all of the sites. It was particularly difficult to recruit homogenous groups of doctors and managers. Following discussion, we agreed that where sites were unable to assemble homogenous groups, site contacts should select participants from pre-existing groups within clinical areas. This compromise was considered acceptable because the original decision to segment participants into homogenous groups reflected a desire to encourage within-group interaction. In all, nine focus groups ran between December 2000 and February 2001 (Table 5.3.2.2).

(e) Running the groups

Group homogeneity, facilitator skill and an appropriate environment are all elements in the strategy for eliciting views, and there is some room for pragmatism in trade-offs between these elements (Stewart and Shamdasani 1990). The practical difficulties involved in bringing together participants is acknowledged by Morgan and Krueger (1993), who cite geographically dispersed populations as an example. While the present study populations were geographically concentrated, they consisted of groups of professionals working to capacity and with limited time for extraneous activities.

Table 5.3.2.2: Composition of focus groups, by site

Site A: Community and Mental Health Trust [PHC]	Group 1	x6 Community Nurses x3 Clinical Psychologists
	Group 2	x6 Health Visitors
Site B: Acute Hospital [GHH]	Group 1	x3 Nurses x1 PAM x1 Midwife x2 Operational manager
	Group 2	x1 Midwife x2 PAMs x2 Nurses x1 Nurse manager
Site C: Acute Hospital [MSGH]	Group 1:	x1 Consultant Surgeon x1 Ward Sister x2 Nurse Managers x1 Orthoptist x1 Health Care Support Worker x1 Anaesthetist
	Group 2:	x3 Consultants x4 Managers x1 Nurse
Site D: Acute Hospital [GEH]	Group 1: Urology	x1 Consultant x5 Nurses x2 Community Nurses
	Group 2: Trauma	x1 Consultant x1 Junior Surgeon x3 Nurses x1 OT x1 Physiotherapist
	Group 3: Audiology	x1 Consultants x4 Nurses x1 Nurse Manager x1 Junior Doctor

The second recruitment strategy of selecting participants from specific clinical areas placed greater responsibility on the moderator to secure participation within the groups, to offset the professional heterogeneity of the groups. Under these circumstances, the moderator

deals with the potential inhibiting factors due to group composition. Of particular importance in this regard are hierarchical distinctions between group members, which require careful management to encourage lower status participants to contribute. This was secured principally by verbal and non-verbal signals in communication, identifying when participants had contributions to make in order to give them the floor. Further, participants were provided with detailed information during the introduction, and explored ground rules for listening and talking before each session. The ground rules explicitly addressed the importance of valuing contributions from all participants, and the need for active and reflective listening.

(f) Tape recording and transcription

Both analytical and pragmatic considerations influenced the decision to tape record the focus groups. The former was due to the rich access accorded to the discussions (Barbour & Kitzinger 1999), the latter to the absence of a co-facilitator to make contemporaneous notes. The introduction to each focus group included a statement of the importance of taping the sessions, stressing the importance of participants' contributions. All participants readily accepted the need for taping, and all of the focus groups were tape-recorded.

Commentators typically recommend full transcription of focus group tape recordings, detailing every half-finished sentence, hesitation and silence (Stewart & Shamdasani 1990), a procedure acknowledged as extremely time consuming. This is entirely appropriate where subsequent analysis requires such detail, as in conversation analysis approaches, which have their own notation systems for pauses, overlaps and changes of turn (Atkinson

& Heritage 1984). Such detailed notation was not required in this application, and consequently the moderator listened to each tape, writing verbatim notes of the discussions but excluding 'ums', errs and pauses, then transcribed the notes into a word-processing package. These word-processed documents formed the data used in subsequent analysis.

(g) Analytical strategies

Compared with the volume of literature on conducting focus groups, there is a relative lack of literature on analysing focus group data. Researchers have used focus groups to collect data for analysis within a wide range of methodological traditions, including among others phenomenology, grounded theory and content analysis (Carey 1994). Some of these applications, particularly within phenomenology and grounded theory, are subject to criticism (Webb & Kevern 2001). However, there is a general acceptance that as with other methods, the research questions and study purposes best determine the nature of data analysis.

Stewart & Shamdasani (1990) outline a typology of different forms of content analysis, including pragmatic, semantic and sign vehicle analysis, each having specific uses. Where the primary purpose is to summarise large amounts of qualitative data by theme, as in the present study, they recommend a simple 'cut and paste' technique, also described by Knodel (1993). The technique has two stages. The former data making stage involves organisation of data into meaningful segments relevant to the research question(s). Once this is completed, the analyst builds a taxonomy of major topics and / or issues by reviewing the

data in each segment. The analysis ends with a summary report, listing the topics and issues and providing examples of the range of comments in each.

As the primary purpose of the research was to identify potential indicators of progress, the simple cut and paste technique was used to organise the data segments into sub-themes within each domain. Once the data was organised in this way, the statements became the basis of indicator items within each sub-theme, covering the range of participants' contributions.

(h) Data handling and preparation

Data handling decisions reflected the primary purpose of the study and the chosen analytical framework. The primary purpose was to identify potential indicators of progress in clinical governance, by exploring staff descriptions of ideal practice and current shortfalls in specific domains of interest. This required an analysis of the positive and negative statements made by participants in each domain, revealing the diversity of comments and capturing the language and vocabulary of participants.

Data making

Also known as data assembly or data reduction, this stage involves the selection, abstraction and simplification of raw data (Miles & Huberman 1994). To facilitate analysis of all statements made about a domain across all focus groups, segments of the transcriptions relating to each domain in the topic guide were assigned an identifying code. This was done with the aid of 'The Ethnograph'; a computer programme designed to

support data handling in qualitative analysis. Use of this programme facilitated speedy and accurate retrieval of all of the pre-coded segments relating to a particular domain, ensuring that all were included in the analysis. Table 5.3.2.3 identifies the number of data segments within each domain.

Table 5.3.2.3: Number of data segments, by domain

Domain	Ideal (n)	Real (n)	Total (n)
1. Adverse events	112	66	178
2. Quality improvement	60	38	98
3. Risk management	49	44	93
4. Learning about practice	57	35	92
5. Organisational change	59	56	115
6. Appraisal	80	38	118
7. Evidence Based Practice	62	36	98
8. Multidisciplinary teams	54	26	80
9. Professional development	15	6	21
Totals (n)	548	345	893

Coding

This is essentially a process of categorising the data in order to facilitate an understanding of its meaning. The categories came from the data itself, generated inductively using an iterative approach. Initially, all segments relating to a single domain were identified by their pre-codes, retrieved using 'The Ethnograph', and printed out to form two assembled files (one ideal practice, the other current shortfall), consisting of relevant data segments. The first third of the data in each file was read closely several times, and following reflection a number of common emergent themes (categories) within the data were noted to

form a provisional coding scheme. The process was repeated twice more with successive thirds of the segments until no data remained, the emergent categories refined and developed further between each successive iteration. This process was repeated separately for each domain, providing a thematic overview of the statements made within each domain based on emergent categories.

5.3.3 Results

Analysis proceeded around 893 data segments. Table 5.3.3.1 provides a summary overview of the sub-themes inductively generated in each domain. There follows a description of each sub-theme, supported by example data segment, chosen to reflect the range of segments contained in the theme. Finally, the section provides an initial pool of CGCQ indicator items based on the original data segments.

Table 5.3.3.1: Domains and their inductively generated sub-themes

Domain	Sub-theme
Adverse events	Openness and encouragement to report Blame Using a broad approach in investigations Making changes where necessary
Quality improvement	System-wide orientation Stakeholder involvement Multi-disciplinary focus Investment in improvement Innovatory environment Quality improvement techniques Planned improvement
Risk management	Risk assessment A corporate approach Action on identified risks
Organisational learning	Investment in learning Sharing learning with others
Organisational change	Motivation to change Support Leadership
Evidence Based Practice (EBP)	Technical skills Access to research evidence
Multidisciplinary team	Common goals Mutual respect Clear roles Participation
Appraisal	Reviews used developmentally
Professional development	Training and development needs Using new skills

Domain 1: Adverse events

(a) Openness and encouragement to report adverse events (Table 5.3.3.1.1)

Given the importance of responding to adverse events, a climate congruent with good clinical governance is one in which clinicians report their involvement in adverse events. The negative corollary is that they feel that there is little point in honest reporting, or worse, that they will be seen as troublemakers.

Table 5.3.3.1.1

Adverse events sub-theme 1: Openness and encouragement to report adverse events	Ideal or Real
1. "There should be a non-judgemental atmosphere to allow reporting; to be encouraged to admit making mistakes and being open enough to apologise." [gehurol:26-29]	Ideal
2. "Openness should be encouraged." [gehtrau: 49]	Ideal
3. "You should be patted on the back for coming forward with something rather than trying to hush it up." [premn:46-47]	Ideal
4. "We would be open about the event." [mscons: 17-18]	Ideal
5. "We would discuss what happened, why it happened and the team would give support, emotional support, getting people to act as a team- if we are a team, we should be there to help one another." [mssurg: 31-36]	Ideal
6. "People that identify adverse incidents are seen as people who are creating problems- they then think 'why bother' next time." [ghh2:86-89]	Real
7. I'm not sure that if things were recorded that management wouldn't use it as a stick to beat us with [premn: 78-80]	Real
8. "Only a fool would admit to making an error." [premn: 96]	Real
9. "Often people can't learn from experiences because people close in - there is only hearsay after an incident. This prevents the support and learning for everybody." [premmh:131-135]	Real
10. "If I made a mistake, I would accept the responsibility of that, to discuss what had happened." [gehchol: 23-25]	Real
11. "If I do something wrong then I'm not going to tell anybody - I'll just keep it quiet." [gehchol: 514-516]	Real

(b) *Blame* (Table 5.3.3.1.2)

This theme refers explicitly to an expectation of unjust blame and punishment of individuals involved in adverse events, rather than support. It was often contrasted to a systems approach to adverse events in which latent systemic causes are identified and corrected.

Table 5.3.3.1.2

Adverse events sub-theme 2: Blame	Ideal or Real
1. "People feeling that they will not be unjustly blamed- allowing them to be open and honest." [ghh1:62-64]	Ideal
2. "There needs to be a blameless culture within the trust, to enable people to be honest." [ghh2: 17-19]	Ideal
3. "People who are involved should receive support through the process- not made to feel guilty until proven innocent." [ghh2: 45-48]	Ideal
4. "We should avoid thinking automatically that there is someone to blame." [mscons: 91-93]	Ideal
5. "Underlying causes being looked for and changes made to the underling causes, rather than individuals blamed." [premn: 55-58]	Ideal
6. "You would feel able to report the incident without feeling that the whole world will be let loose on you - and you would lose your job." [ghh2: 21-24]	Ideal
7. "The person who has made the mistake is made to feel like a criminal- they feel like 'everyone is out to nail me'." [gehurol: 64-67]	Real
8. "Once there has been an event, those at the centre are watched closely and not trusted again- they are tainted by the experience." [gehurol: 116-119]	Real
9. "People look for a scapegoat- who made the mistake rather than how did the mistake occur?" [ghh1: 76-78]	Real
10. "There is a 'there but for the grace of God go I' approach - those that don't understand the realities of working in a clinical area might be more critical in assessing a situation. So the need to blame is tempered by the need not to blame also." [ghh1: 101-110]	Real

(c) *Using a broad approach in investigation* (Table 5.3.3.1.3)

This refers to the use of systems approaches to the investigation of adverse incidents, in which an examination is made of latent problems within systems of care. Events are placed within their broader organisational context, in order to avoid unjust scapegoating of individuals.

Table 5.3.3.1.3

Adverse events sub-theme 3: Using a broad approach in investigation	Ideal or Real
1. "There would be an examination of what went wrong, how it went wrong to find the true cause, whether an individual or a system." [gehtrau:21-24]	Ideal
2. "The investigation should be opened up widely; you can get blinkered too quickly looking for cause and effect- almost knee-jerk reactions happen." [gehtrau:51-54]	Ideal
3. "You need to examine which underlying processes have failed, not just blame the individual at the centre." [gehurol: 41:44]	Ideal
4. "The whole process and circumstances should be examined, the activity at the time etc., instead of highlighting the error, the total situation should be examined." [ghh2: 34-38]	Ideal
5. " We need to relate 'problems' to the wider context in which the case takes place without blaming individuals- look at system failure rather than blame the person." [premn: 35-37]	Ideal
6. "I think one of our problems is the depth to which we explore an issue. (...) I wonder whether we really do sit down and look beyond initial problems. I don't think we do any more than scratch the surface." [gehtrau: 90-99]	Real
7. "We often get a reaction that is knee-jerk and there is a need to 'solve' things straight away. We just change the way we do things and are still doing it badly!" [gehurol 97:101]	Real
8. "Things happen as a result of situations, and we have to work out how the situation can be avoided in the future. It might be due to overwork, it might be due to stress, pressure etc." [gehchol: 41-46]	Ideal

(d) *Making changes where necessary* (Table 5.3.3.1.4)

This refers to an active response to adverse events, an attempt to avoid repetition of a negative event, by implementing those changes to practices and systems suggested by a thorough investigation of 'what went wrong'.

Table 5.3.3.1.4

Adverse events sub-theme 4: Changes where necessary	Ideal or Real
1. "Circumstances examined, staffing load at the time was assessed, policy & procedures reviewed." [gehurol: 78-80]	Ideal
2. "You should be putting something in place to ensure that it doesn't happen again." [ghh1: 34-36]	Ideal
3. "Action should be aimed at making systems changes that avoid the same conditions arising again in the future." [ghh2: 50-53]	Ideal
4. "We would review procedures and practice to reduce the chances of the error occurring again." [mscons: 13-15]	Ideal
5. "We have to be sure that we are not just reporting; that we are also acting on our reports." [mscons: 35-37]	Ideal
6. "We would need to identify implications and act; either by informal processes, or more formal processes of reviewing incidents." [premmh: 86-89]	Ideal
7. "We never change practice as a result of an investigation into adverse incidents." [gehtrau: 132-134]	Real
8. "Systemic problems are taken into consideration, but they are not addressed - there is no learning programme leading to systemic change. There is a gloss - a desire that things will never happen again; but practices quickly fall back to the way they were - changes are not sustained." [ghh2: 65-73]	Real

Domain 2: Quality improvement

(a) *System-wide orientation* (Table 5.3.3.2.1)

Participants identified the importance of a broad focus for quality improvement initiatives rather than on simple 'within-team' fixes.

Table 5.3.3.2.1

Quality improvement sub-theme 1: System-wide orientation	Ideal or Real
1. "... after there is a review in a single team which highlights that other parts of the service need changes - we often don't have the power to change things in the wider system which would have been useful. There is a general sense of 'go away'. The wider context of change is important."	Ideal
2. "We need to look at the effects of change in one part of the system on the wider system." [mscons: 155-157]	Ideal
3. "We need to look at broad systems of care and make changes that get at the root cause of problems - working across teams to alter the systems." [mscons: 158-162]	Ideal
4. "Solutions to quality improvement, risk management and adverse events need to take account of the potential impact on other teams. There is no point solving one problem and creating 3 others elsewhere. The wider effects need to be taken into consideration." [gehtrau: 233-240]	Ideal
5. "Often the problem we are trying to solve exist somewhere else - we try and address the problem that requires major inputs from elsewhere, that we cannot address." [premmh: 259-263]	Real
6. "We tend to be insular and try to improve our area rather than act trust-wide." [mssurg: 238-240]	Real
7. "We find out about changes in other parts of the system only when their effects are felt - we are not included in discussions about things that affect us." [ghh1: 291-295]	Real
8. "Sometimes changes happen before we know anything about it and we just have to pick up the pieces - we are not involved in the process, and then held to account for doing things the 'new' way." [premmh: 476-481]	Real
9. "Often changes are not discussed before being implemented." [gehurol: 310-313]	Real

(b) Stakeholder involvement (Table 5.3.3.2.2)

This sub-theme identified the importance of stakeholder perspectives in quality improvement, including those of patients, managers and clinicians.

Table 5.3.3.2.2

Quality improvement sub-theme 2: Stakeholder involvement	Ideal or Real
1. "Patients would be involved in service change, their perspectives are important." [ghh2: 109-111]	Ideal
2. "There would be input into the process from customers - whoever there might be: nursing staff, patients, whoever they might be." [mscons: 142-145]	Ideal
3. "Explicit expectations between all stakeholders, and a recognition of the reasons between different stakeholders' expectations. Even in ideal worlds, managers, users and clinicians will want different things. So we need to know what those differences are and reach some agreement on how to proceed." [premmh: 198-206]	Ideal
4. "Stakeholders pull in different directions - people hang on to their own perspective which stops progress forward. Stakeholders often don't communicate well with each other - a divide between what different stakeholders want." [premmh: 235-241]	Real
5. "Often we are focussed on what the service needs rather than what the clients need. The vision of a high quality service gets lost in what we need to be happy with our part of the service." [premmh: 243-248]	Real

(c) *Multi-disciplinary focus* (Table 5.3.3.2.3)

Participants identified the importance looking beyond the immediate environment to include multidisciplinary teams, others from different teams, and across different sectors of the health care economy.

Table 5.3.3.2.3

Quality improvement sub-theme 3: Multi-disciplinary focus	Ideal or Real
1. "Change would be multidisciplinary - not everyone in their own little world, but involve people more and have a shared vision." [ghh2: 93-96]	Ideal
2. "Dialogue between clinical and directorate teams is important; so that concerns can be raised and there are clear responses." [gehtrau:159-162]	Ideal
3. "We would be involving other people from across primary and secondary care and whoever has an input into care - wider working relationships so that we are not simply working in isolation. Ideally this would be seen as relevant - you would not be seen as not doing what you are employed to do if you do this - it would be seen as relevant." [gehtrau: 504-513]	Ideal
4. "Working across teams is not something that we routinely do - we tend to work within our own area." [mscons: 181-183]	Real
5. "There are blocks between directorates - everyone seems to be fighting their own corner rather than looking at the bigger picture." [mscons: 227-230]	Real

(d) *Investment in improvement* (Table 5.3.3.2.4)

Staff believe that the real importance of quality improvement to the organisation is indicated by the willingness or otherwise to provide resources to facilitate improvement, principally time. A climate conducive to good clinical governance would ensure that immediate pressures would not override quality improvement activity. The reality for many was that immediate pressures were more important.

Table 5.3.3.2.4

Quality improvement sub-theme 4: Investment in improvement	Ideal or Real
1. "Quality improvement would be seen as an important activity and time would be prioritised to allow clinicians to do it." [ghh1: 120-123]	Ideal
2. "Teams would have time to reflect rather than simply be reactive all the time." [mscons: 130-132]	Real
3. "I do think that if we're honest there is an element of fire-fighting in what we do. You tend to sort out today's problems. It would be nice to step back and make real improvements. All of us would like to be able to step back and say 'this is what we would like to do to improve quality' but it gets lost in the day to day." [gehtrau: 202-211]	Real
4. "Some people are very possessive of what they do and cannot see the benefit of any change." [gehurol: 171-173]	Real
5. "There is no protected time for me to become involved in quality improvement activities - other things will override this is priority." [ghh1: 152-155]	Real
6. "Quality improvement is not seen as a high priority - immediate pressures are always more important." [mscons: 185-189]	Real

(e) *Innovatory environment* (Table 5.3.3.2.5)

This refers to the support given to allow autonomous practitioners to take a lead in quality improvement; the sense that it is under their own control, with the support and provision to make it a reality.

Table 5.3.3.2.5

Quality improvement sub-theme 5: Innovatory environment	Ideal or Real
1. "People would accept that quality improvement is an ongoing thing - we can't just assume that we have got it and it will stay that way forever. It is an ongoing thing to maintain the standards that you have achieved." [gehtrau: 170-175]	Ideal
2. "An acceptance of the notion that change for the better is part of the job." [gehtrau: 138-140]	Ideal
3. "Change should be encouraged from the bottom up, not imposed from the top down. People would feel involved in the change - they would be changing not being changed." [ghh2: 98-104]	Ideal
4. "People are simply ascribed the task of quality improvement rather than cherishing and wanting it - 'you have got to improve this' from above rather than because they are interested." [gehurol: 165-169]	Real
5. "Quality improvements are often imposed from above - it is often done to us rather than us doing it. The vision is imposed." [ghh2: 113-116]	Real
6. "We often feel coerced into making change rather than agreeing with the need for change and embracing it." [gehurol: 298-300]	Real
7. "Change is externally imposed - we are told to change. We were not asked for our opinion so it felt scary. There was no facility to observe or practice a change in delivery - just told to do it, with no follow-up." [premmh: 446-452]	Real

(f) Use of quality improvement techniques (Table 5.3.3.2.6)

Participants identified the importance of a diagnostic phase to explore current shortfalls in care quality and identify opportunities for improvement. There was little discussion of specific quality improvement techniques, such as control charts. Comments were limited to clinical audit technologies. There was no mention of other options, such as use of control charts.

Table 5.3.3.2.6

Quality improvement sub-theme 6: Use of quality improvement techniques	Ideal or Real
1. "Setting standards based on good practice and auditing practice against this. And collecting information on what is currently happening and examine this against a notion of good practice." [gehurol: 123-128]	Ideal
2. "Auditing against standards." [ghh1: 128]	Ideal
3. "Looking for gaps between what we are striving for and what we are doing as a team. Also looking at patient perspectives of care." [mssurg: 99-102]	Ideal
4. "Sometimes what you think you need is not necessarily what you do need - we need a process to identify these." [premmh: 193-196]	Ideal
5. "There are no formal quality improvement processes - it is not co-ordinated." [gehurol: 161-163]	Real

(g) *Planned improvement* (Table 5.3.3.2.7)

This theme relates to clarity about the purpose of planned quality improvement, as well as the planning process itself. It was considered important to conduct activity in the light of reflection on its aims and objectives, and that such activity was rationally defensible rather than simply activity for its own sake, to be seen to be 'doing something'. Further references to bodies of empirical evidence underlined the importance of rationality within this decision-making process.

Table 5.3.3.2.7

Quality improvement sub-theme 7: Planned improvement	Ideal or Real
1. "You need a clear idea of the quality improvement objective - unless you are clear you can end up with a 'scatter-gun' approach and miss the target" [gehtrau: 153-157]	Ideal
2. "They need to make changes - in order or anything to happen as a result of searching, they need to talk and make decisions about what they are going to develop - so plan and develop" [mssurg: 374-378]	Ideal
3. "You need to know what your aims and objectives are for improvement" [mscons: 134-135]	Ideal
4. "Change that is appropriate for quality improvement is seen as rational - not change for its own sake" [gehtrau: 178-181]	Ideal
5. "If we want to change practice, we have to make a case for it with reference to the evidence. If you want to make a change you should present the evidence for it" [gehtrau: 529-533]	Real
6. "Much of our 'quality improvement' is a knee-jerk response to a problem without thinking it through" [gehurol: 156-159]	Real

Domain 3: Risk management

(a) *Assessment* (Table 5.3.3.3.1)

There is a basic requirement for assessment procedures as the basis of planning.

Assessment processes require staff awareness of inherent risks, and an understanding of their importance. Participants revealed that assessments were often ad hoc and intuitive, rather than systematically planned.

Table 5.3.3.3.1

Risk management sub-theme 1: Assessment	Ideal or Real
1. "There should be a systematic way of assessing for risk and preventing problems. It is done largely as an intuitive thing. This doesn't feel right, there should be more than this" [gehtrau: 242-247]	Ideal
2. "Risks need to be assessed. Time should be protected to identify and plan" [gehtrau: 249-253]	Ideal
3. "Everyone would be aware of the clinical risks involved in their work and understand why it needs to be monitored" [ghh1: 163-166]	Ideal
4. "The system for collating information on risks is not widely used" [ghh1:211-213]	Real
5. "We do risk assessments and see them as part of our job" [ghh1: 227-231]	Real
6. "We don't do formal assessments- we don't have formal systems" [mssurg: 176-177]	Real

(b) *A corporate approach* (Table 5.3.3.3.2)

This theme refers to the felt need for a common, organisation-wide framework for identifying and managing clinical risks. Such a system ensures the management of risks common to many areas, such as medication errors or falls, in the safest way and in a similar fashion across the organisation. Other important elements include prioritising risks, in order to expend effort on the most important areas. These are likely to include areas where error could result in catastrophic consequences for patients, such as serious injury or death. This relies on proactive localised activity to derive, share and update protocols.

Table 5.3.3.3.2

Risk management sub-theme 2: A corporate approach	Ideal or Real
1. "There is a risk group that co-ordinates information across the trust" [gehtrau: 215-217]	Ideal
2. "There are protocols and guidelines for specific categories of risk. Protocols would be widely shared across the hospital / trust" [gehtrau: 222-226]	Ideal
3. "Co-ordination of risk management strategies - so there is a common policy on similar risks across the whole organisation. Some central generic assessment, then locally implemented - so common standards across the trust. There is no corporate vision for sharing risk management standards across the organisation" [ghh1: 171-182]	Ideal
4. "There would be an assessment of the relative risks of all major problems facing the trust - a helicopter view and prioritising process for managing risk" [mscons: 206-210]	Ideal
5. "We need a comprehensive system to identify what the risks are - then need to capture where it is felt that there have been problems or potential problems, that brings together different elements" [premmh: 288-293]	Ideal
6. "Things are piecemeal here at the moment - we do not share them across areas so common problems may be handled differently. Risk management is not proactive. We don't look at risk systematically to try and find loopholes and weaknesses. We just do it in a patchy way and almost wait for things to go wrong. There is no common interrogation of risk management data or learning across directorates / departments based on how risks are managed - 'splendid isolation'" [gehtrau: 255-271]	Real
7. "There is no central collation of risk management data looking for trends or sharing protocols" [gehtrau: 288-290]	Real

(c) *Action on identified risks* (Table 5.3.3.3.3)

Following assessment, participants identified a need for action to deal with the identified risks; failure to act renders the process null and void. Participants identified the negative impact on motivation of failure to act on identified risks. Such inaction promoted scepticism, leading to a reassessment of risk management activity as a 'paper exercise'.

Table 5.3.3.3.3

Risk management sub-theme 3: Action on identified risks	Ideal or Real
1. "There would be an appropriate response to identified risks - documented action taken in the light of the assessment. Unless assessments are acted on they are worthless" [gehtrau: 249-253]	Ideal
2. "Recommendations based on assessment of risks are pushed through and have real effect" [mscons: 191-196]	Ideal
3. "Every time there is a change in the process, system or technology, risks need to be reassessed" [mssurg:140-142]	Ideal
4. "Risk management is seen largely as a paper exercise - it doesn't lead to actions on identified risks" [gehtrau: 280-282]	Real
5. "There is no urgency to deal with potential issues" [ghh1: 215-216]	Real
6. "Sometimes risks that have been identified are simply not addressed" [mscons: 198-199]	Real

Domain 4: Organisational learning

(a) *Investment in learning* (Table 5.3.3.4.1)

This sub-theme partially echoes that of sub-theme 5 of quality improvement, 'environment for improvement', as it includes elements of a supportive external environment that facilitates reflection on practice for improvement, by public recognition and provision of resources such as time. However, it also contains an internal attitudinal dimension, in which practitioners subscribe to the importance of their own continuous learning, and accept as legitimate the learning needs of others.

Table 5.3.3.4.1

Organisational learning sub-theme 1: Investment in learning	Ideal or Real
1. "It would be seen as important - sharing quality improvement would be seen as an essential part of best practice and quality improvement" [mscons: 230-233]	Ideal
2. "The feeling that it is OK to spend time on such activities - to attend fora where ideas are shared" [gehurol: 227-229]	Ideal
3. "People able to feel that it is acceptable to devote time to dissemination" [ghh1: 238-40]	Ideal
4. "There would be protected time available to people to reflect on their practice" [ghh2:166-168]	Ideal
5. "There would be a supportive environment to share insights between teams and different areas of the trust" [ghh2: 170-173]	Ideal
6. "People would spend time in other departments to find out different perspectives" [mscons: 235-237]	Ideal
7. "The reality is that practice development is done in our own time - the reality is that you come in to work. If you spend time on development, people will question you - 'what are you doing' - so you have to do it in your own time" [gehtrau: 343-349]	Real

(b) *Sharing learning with others* (Table 5.3.3.4.2)

Many participants expressed a sense of geographical and professional isolation, yet were able to identify a range of dimensions to learning from others, both within and beyond the organisation. They noted that maximum benefit from learning required openness between staff; an emphasis on co-operation for improvement rather than competition between teams and individuals and an associated need for secrecy.

Real openness includes the sharing of practice that was 'not so good', with the intention of learning how to improve. Such vulnerability may only be shown when staff feel that it is safe to admit their present lack of perfection. The reality for some was that performance other than exemplary is considered shameful and needs to be hidden away. In such a climate, the sharing of problems in order to identify potentially helpful solutions is impossible

Although participants considered trust-wide networks useful in fostering improved learning opportunities within organisations, the importance of working beyond the organisation was also emphasised. Regional and national networks were characterised as important sources of information on best practice. Such networks were considered useful in avoiding organisational insularity.

Table 5.3.3.4.2

Organisational learning sub-theme 2: Sharing learning with others	Ideal or Real
1. "There are opportunities for learning from others across the organisation and beyond it. There need to be forums for individuals to feed back to the trust as a whole, not just within their directorates" [gehtrau: 315-322]	Ideal
2. "There would be an openness to sharing practice that wasn't so good - not just showing the best all the time" [gehtrau: 328-332]	Ideal
3. "People would work together to improve services, not be in competition with each other. There would be no need to be secretive" [ghh2:183-186]	Ideal
4. "We would learn how people in other organisations perform their work; it is so easy to get insular in how you do things" [mscons: 242-245]	Ideal
5. "Benchmarking against other organisations would take place" [gehtrau: 491-492]	Ideal
6. "There would be good networking with national and regional special interest groups for exchange of information on good practice" [gehtrau: 499-502]	Ideal
7. "There are some opportunities for sharing, but most people in the trust are unaware of good practice in other teams / directorates. We have people who are leading the way in care regionally, but do not know any details of their work" [gehtrau: 339-341]	Real
8. "We need to link across the organisation - people are encouraged to work geographically and professionally in isolation so what happens in one area of the trust doesn't happen in another, regardless of whether it is good practice. So the service is quite different in different areas" [premmh: 370-378]	Real
9. "We only hear about different teams when there has been a negative incident or a complaint - we try and hide these. Some people don't share anything - things happen and it is swept under the carpet. The opportunities to share are lost" [premn: 210-218]	Real

Domain 5: Organisational change

(a) *Motivation to change* (Table 5.3.3.5.1)

This sub-theme refers to the willingness of staff to reflectively question their practice, and try new approaches to service delivery where warranted.

Table 5.3.3.5.1

Organisational change Sub-theme 1: Motivation to change	Ideal or Real
1. "Managers and peers would be open-minded and prepared to try new approaches. People would be prepared to take a measured risk in changing things - not accept that this is the way it's always been so this is how it will remain" [ghh2:208-215]	Ideal
2. "When people agree to make a change they follow it through: so often when protocols are agreed at departmental meetings, people just won't work in that way and the whole thing collapses" [mscons: 312-317]	Ideal
3. "Sometimes we get into a bit of a loop - we do things because we have always done them that way. I think it is quite important that it is OK to question the way that we do things" [gehtrau: 74-79]	Ideal
4. "The ability to question custom and practice - you can be so close to something that you don't actually see anything" [mscons: 451-454]	Ideal
5. "There is a great deal of inertia in the trust which stops most change ... there is a lack of willingness to do things differently to how they were done in the past" [gehtrau: 388-393]	Real
6. "People bury their heads: they think 'we've always done it this way and it's worked up 'til now - why change?'" [ghh1: 317-322]	Real
7. "People can dig their heels in and refuse to accept any change - they see it as potentially disrupting their lives and oppose any change" [ghh2: 225-230]	Real
8. "Even simple little changes do not happen here - the level of inertia is too high" [mscons:173-175]	Real
9. "Changes seem to be reactive and crisis led" [premmh: 468-469]	Real

(b) *Support* (Table 5.3.3.5.2)

This is partly related to quality improvement sub-theme 5 (environment for improvement), in its emphasis on a climate that encourages clinicians to make the changes that they deem necessary, and enable them to innovate without fear of punishment. This is combined with the practical resources required, such as time and training to develop skills.

Table 5.3.3.5.2

Organisational change sub-theme 2:	Support	Ideal or Real
1.	"There needs to be some organisational vision - not just lip service to the ideal of clinical governance but real investment for future benefits" [gehtrau: 366-370]	Ideal
2.	"Safety - not to feel threatened because of lack of information but to be supported through the change process" [gehurol: 276-279]	Ideal
3.	"People would have change management skills - not be expected to make changes without having appropriate changes" [ghh1: 286-289]	Ideal
4.	"People are less likely to try to push the boundaries when there is a fear of failure - a hell of a fear of getting it wrong. It is very tempting to play safe and we see that in a lot of our practice. People are frightened of making mistakes" [gehtrau:395-403]	Real
5.	"I find that the difficulties can come from above. I work well with my staff to make practice changes, but there can be difficulties. To take clinical care forward can be difficult. I don't get stopped, but I do get some queries. It is not normal here to be continually changing" [mssurg:282-289]	Real
6.	"We ask people to change without giving them training to cope with that change so people say 'I'm not doing that' " [premmh:454-457]	Real
7.	"Often we identify areas to develop but have no financial resources to then do it" [premn: 251-253]	Real

(c) *Leadership* (Table 5.3.3.5.3)

This refers to the nature of relationships between managers and clinicians and the active involvement of all staff in leadership activity. Ideally, managers would be open to innovative ideas from clinicians, and be supportive in developing plans for their realisation. The subjective experience for staff would be the welcoming of ideas, rather than their dismissal as a distraction.

Many participants conceptualised leadership as the discrete function of leaders, rather than a responsibility of all. Thus, the responsibility for action on a good idea is considered to rest with others.

Table 5.3.3.5.3

Organisational change sub-theme 3: Leadership	Ideal or Real
1. "Management would be receptive to ideas for change" [gehurol: 257-258]	Ideal
2. "You would know that if you had a good idea it would be listened to" [mscons: 282-283]	Ideal
3. "If I had a good idea I'd know where to go to get it moved forward" [mscons: 285-286]	Ideal
4. "The expectation that if I make a suggestion for change that everyone agrees is right then it will happen" [ghh1: 282-284]	Ideal

Domain 6: Evidence-based practice (EBP)

(a) *Technical skills* (Table 5.3.3.6.1)

Participants identified a range of technical skills required by clinical team members, including search strategies, critical appraisal, and the ability to then build coalitions for change within teams, to implement and review practice improvements.

Table 5.3.3.6.1

EBP sub-theme 1: Technical skills	Ideal or Real
1. "There would be appraisal skills to allow you to identify good practice" [gehtrau: 483-485]	Ideal
2. "The skills to search literature and find information on the best practice" [premmh: 603-605]	Ideal
3. "With a team, access to appraisal and searching skills - not one expert, but a range of people coming together" [premmh: 608-611]	Ideal
4. "Clinical audit infrastructures would be available to help develop EBP - the technical aspects of auditing" [gehtrau:494-497]	Ideal
5. "Skills of appraisal - technical appraisal, and communication skills - need to be able to discuss with the team what is required and then agree a programme of work - not just develop it in isolation" [gehurol: 392-398]	Ideal
6. "People have the skills to take EBP forward - to translate all the reviewing activity into real practice change, through guidelines for example" [ghh1: 381-385]	Ideal
7. "Technical skills in appraising the quality of research - critiquing the papers identified" [ghh2: 285-287]	Ideal
8. "I don't know who to approach for information or advice on critical appraisal of research" [gehtrau: 547-551]	Real
9. "We don't utilise the EBP skills that we have in the team. We have no way of knowing who has these skills within the trust - it is not collated" [ghh2: 312-319]	Real
10. "We have internet access, but we have only limited search skills. I rely on colleagues." [premmh: 648-650]	Real

(b) *Access to research evidence* (Table 5.3.3.6.2)

Skills in searching for research evidence and appraisal of the quality of retrieved research are of limited value unless clinical teams have access to sources of information, both paper and electronic. Specific resources included Internet access to electronic research databases such as Medline and Cochrane. The local availability of full-text articles was important, as abstracts on their own were of limited value.

Table 5.3.3.6.2

EBP sub-theme 2: Access to research evidence	Ideal or Real
1. "Access to the internet and other electronic sources of information, as well as good internal information systems" [gehtrau: 148-151]	Ideal
2. "There would be good access to internal information from patient administration systems" [gehtrau: 487-489]	Ideal
3. "People would have access to the internet" [ghh1: 375]	Ideal
4. "Training opportunities are available to support technical skills and implementation" [ghh2: 308-310]	Ideal
5. "Access to the internet - Medline and Cochrane. Access to electronic and paper sources" [mssurg: 365-367]	Ideal
6. "Access to full text literature rather than just abstracts, available locally" [Premn: 354-356]	Ideal
7. "There is a frustration with not being able to access what information there is" [mscons: 460-462]	Real
8. "I didn't know that I had access to libraries and electronic facilities!" [mssurg: 409-410]	Real
9. "Often access is unequal between different members - Doctors get better access" [premn: 347-349]	Real

Domain 7: Multidisciplinary team

(a) *Common goals* (Table 5.3.3.7.1)

Respondents identified the importance of a clear, shared understanding of the goals and aspirations of clinical team members. The experience would ideally be one of co-operation rather than competition, in which clinicians' felt that they were part of a collective enterprise, sharing their colleagues' aims.

Table 5.3.3.7.1

Multidisciplinary team sub-theme 1: Common goals	Ideal or Real
1. "There is a common team vision - we are all pulling in the same direction. We know what our priorities are. Everyone is committed to the goals of the team" [gehtrau: 570-575]	Ideal
2. "There are shared goals" [mscons: 475]	Ideal
3. "We work well together - we are all singing from the same hymn sheet" [gehtrau: 601-602]	Ideal
4. "People would work together as a team, and this would be valued" [mssurg: 424-425]	Ideal
5. "Common working across organisational boundaries - a common vision and way of working" [premmh: 707-709]	Real
6. "There is no shared ownership of direction within the groups" [gehtrau: 590-591]	
7. "People seem to compete against each other rather than work with each other" [gehtrau: 597-599]	Real
8. "There is not really a clear vision of what the team is trying to achieve ... Our aims and objectives are not openly discussed and agreed" [ghh1; 414-421]	Real
9. "Team members are very competitive and do not work together to provide a common vision" [ghh2: 346-348]	Real
10. "I feel isolated - I don't feel part of a team" [mssurg: 427-428]	

(b) *Mutual respect* (Table 5.3.3.7.2)

Ideally, relationships between members would reflect mutual respect for the professional contribution of each. This would be revealed through open and democratic decision-making, rather than diktat based on formal hierarchical position.

Table 5.3.3.7.2

Multidisciplinary team sub-theme 2: Mutual respect	Ideal or Real
1. "They need to be able to go beyond the hierarchies - each individual needs to develop services. The NHS is historically very hierarchical: people need to go beyond this and act as a team - not dominated by any particular professional group" [mssurg: 384-390]	Ideal
2. "There is mutual respect and value for everyone's contribution" [gehtrau: 553-554]	Ideal
3. "We have broken down the old tribalism between doctors, nurses and managers. We are able to get away from stereotypes of what doctors and nurses can do" [gehtrau: 578-580]	
4. "Everyone has the same standing in the team, regardless of their professional background" [gehurol: 444-447]	Ideal
5. "Mutual respect between all members" [mscons: 470]	Ideal
6. "People would not feel that they were in a hierarchy - people would be able to champion the changes that they thought were necessary" [mssurg: 415-418]	Ideal
7. "Equality of voice between participants, and skills in recognising difference. The society in which we live gives a more powerful voice to some and in reality some people do have more experience in some things. It is then a case of how as a team you manage those differences in a useful way, rather than 'well I'm a doctor so I know best so you lot can just shut up' " [premmh: 676-686]	Ideal
8. "People don't feel that everyone has the same value in teams" [ghh1: 266-267]	Real
9. "I feel that teams are still consultant lead" [mssurg: 430-431]	Real
10. "Even when we identify research evidence, doctors will unilaterally decide what we should do" [ghh2: 296-298]	Real
11. "It feels like warring factions sometimes" [mssurg: 442]	Real

(c) *Clear roles* (Table 5.3.3.7.3)

This relates to clarity of understanding of the professional expertise of others by each team member, as well as clarity in team members' different roles within the team.

Table 5.3.3.7.3

Multidisciplinary team sub-theme 3: Clear roles	Ideal or Real
1. "Members [of teams] need to know what others are trying to achieve - what their contribution is" [gehtrau: 556-558]	Ideal
2. "There is a good knowledge of the skills of each team member" [gehtrau: 567-568]	Ideal
3. "We have a full understanding of the different roles of each team member " [gehtrau:586-588]	Ideal
4. "Team members would have a good understanding of everyone's professional role and role within the team" [ghh1: 409-412]	Ideal
5. "We each have a specific role within the team" [ghh2: 340-341]	Ideal
6. "I am clear about my role in the team" [mscons: 480-481]	Ideal
7. "Often members of teams are excluded from change programmes. There needs to be an understanding and appreciation of what people do. People are not aware of what other people do and what their job is. The bigger picture is lost." [ghh1: 297-304]	Real
8. "There are no clear roles and responsibilities for different participants" [mscons: 490-492]	Real

(d) *Participation* (Table 5.3.3.7.4)

The importance of colleagues being able to participate in decision-making activities was noted. This included the discussion of clinical issues, in which team members were able to have robust discussions, tempered by professionalism. Participants valued challenge from colleagues, but the manner in which challenge was conducted was important. Searching questions from a respectful colleague was considered appropriate; malicious challenge was not.

Table 5.3.3.7.4

Multidisciplinary team sub-theme 4: Participation	Ideal or Real
1. "The recognition that challenging others is OK but part of professionalism" [mscons: 456-458]	Ideal
2. "Everyone should feel able to fully contribute to the work and direction of the team" [gehurol: 435-437]	Ideal
3. "Negotiation and discussion within the team to make decisions" [gehurol: 450-451]	Ideal
4. "All team members are encouraged to be open and honest with each other" [ghh2: 337-338]	Ideal
5. "Team members feel able to raise clinical issues with each other" [ghh2: 343-344]	Real
6. "There is robust discussion and banter between different members - not malicious but searching" [mscons: 494-499]	Real
7. "Everyone would feel able to contribute" [mssurg: 421-422]	Real

Domain 8: Appraisal

(a) *Reviews used developmentally* (Table 5.3.3.8.1)

Comments identified the importance of a shared understanding that reviews were a developmental tool, rather than to be used to 'catch people out'. As such, it should be dynamic and used throughout the year to support development opportunities. This requires a level of trust between participants.

Table 5.3.3.8.1

Appraisal sub-theme 1: Reviews used developmentally	Ideal or Real
1. "Appraisal needs to be done with quality - it should not be a punitive measure or used as a whipping tool" [gehtrau: 422-424]	Ideal
2. "It would be a dynamic document - it is no good doing it then leaving it on the shelf for a year - it should be used all the time and be referred back to" [gehtrau: 426-430]	Ideal
3. "It should be honest and open - a feeling that you can say how you really feel, not have to give the right answer to secure your promotion" [gehurol: 329-333]	Ideal
4. "There would be a shared understanding that appraisal is an opportunity for positive development, not a way of catching people out" [gehurol: 385-388]	Ideal
5. "If you are appraised by your line manager you feel you have to be very careful about what you say" [gehurol: 344-346]	Real
6. "People don't understand what appraisal is about - they are suspicious of it and think that it is a way for management to catch them out or spy on them. They are threatened and don't see it as their time; it is big brother looking down" [gehurol: 373-379]	Real

Domain 9: Professional development

(a) *Training and development needs* (Table 5.3.3.9.1)

Participants valued the opportunity to identify developmental opportunities congruent with the needs of the organisation. While the opportunity to identify opportunities was necessary, it was not sufficient. The delivery of formal or informal development opportunities identified in annual planning was crucial. Similarly, participants valued active support in the process. Simply being told to 'go ahead and deal with this' was deemed inadequate and likely to result in disillusion.

Table 5.3.3.9.1

Professional development sub-theme 1: Training and development needs	Ideal or Real
1. "Education opportunities agreed at annual review are delivered" [mscons: 377-378]	Ideal
2. "Appraisal would be matching people's interests and potential for development with the needs of the organisation - needs to be balanced" [ghh2:252-255]	Ideal
3. "Goals identified for development are not checked against - they are left unaddressed. I am supposed to just go and deal with it without support" [gehtrau: 445-449]	Real
4. "The trust is very positive about appraisal, but in terms of linking individual appraisals to an overall trust development plan this is early days" [mssurg: 347-350]	Real

(b) *Using new skills* (Table 5.3.3.9.2)

Participants identified the importance of being able to use skills learned as part of the staff development process. Without opportunities to implement and practice, new skills soon atrophied, leaving staff demoralised and resentful.

Table 5.3.3.9.2

Professional development sub-theme 2: Using new skills	Ideal or Real
"I've been on a number of courses, but these have not been acknowledged, and I've not been able to put them to good use" [ghh2:239-242]	Real

5.4 Draft indicator item pool based on focus group results

The qualitative data described above was used to generate a pool of item indicators for the CGCQ, in the language used by participants. The statements, reproduced in full in the validation questionnaire reproduced in appendix 4, were constructed according to the prescriptive advice on constructing attitude indicators as described by DeVellis (1991) and Oppenheim (1992) in table 5.4.1 below.

Table 5.4.1: Characteristics of good indicators based on DeVellis (1991) and Oppenheim (1992)

Characteristic of a good indicator	DeVellis (1991)	Oppenheim (1992)
1. Avoid excessive length	√	√
2. Avoid double-barrelled questions	√	√
3. Avoid jargon	√	√
4. Avoid multiple negatives	√	√
5. Avoid ambiguity / words with alternative meanings	√	√
6. Avoid proverbs	-	√
7. Give an option for 'don't know' responses	-	√
8. Beware loaded words such as 'free' or 'natural'	-	√
9. Select items that reflect the scale purpose: a test of the strength of a hypothetical latent variable	√	-
10. Include redundant items in the scale i.e. that explain a similar idea in a different way	√	-
11. Include +ve and -ve worded items to avoid acquiescence bias	√	
12. Avoid a 'stilted' approach- select contentiously worded statements from qualitative pilot work		√

A tick [√] indicates that the author covers includes the characteristic as normative good practice in scale construction

5.5 Content validation of the CGCQ items

This section provides an overview of the validation procedures used in the present study to assess the internal (content) validity of the CGCQ.

5.5.1 Content validation

Reliability and validity are central to any scale development exercise. Put simply, developers need to make sure that scale scores consistently reflect differences in the phenomenon to be measured (reliability), and that they do actually measure what they intend to (validity). This second concept is somewhat vague however, especially since we usually seek to measure abstractions, such as the use of cross-functional teams in quality improvement programmes. How can we know that scales really do measure what they are supposed to measure?

Validation strategies are available for each stage in the scale development process. At the final stage of development, we may empirically test how well scores on the scale correlate with other, concurrently administered and well-validated, measures of the topic ('concurrent' or 'criterion' validity). We may also test whether theoretical assumptions about the abstract construct are confirmed by the measure ('construct validity') e.g. do scale scores differ in expected ways between different sub-populations of respondents? I cover these aspects of validity in section 5.7, which explores the criterion and construct validity of the CGCQ scales in detail.

Yet, validity testing need not be limited to the final stages of scale development. Indeed, there are opportunities during early development for conceptual and empirical approaches to validity testing. Literature reviews can provide conceptual content validity when used to underpin scales, and the qualitative exploration of attitudinal domains provides some empirical support for content validity. In the context of the CGCQ, the focus group study reported in section 5.3 above cover these approaches to validity. At the mid point of scale development, immediately following development of an item pool of indicators, there is a role for *content* validation, in which an expert panel reviews the draft item pool and makes suggestions for improvement.

In order to make the case for content validation clear, it is worth revisiting classical scaling theory, the basis of scaling techniques. Measurement scales consist of a pool of items, created with a specific measurement goal in mind. The items in each scale represent a test of the strength of an underlying latent variable. Theoretically, items are chosen randomly from the universe of all possible statements relating to the construct of interest. While not strictly possible, this serves as a marker of good practice; selected items should aim to exhaust the possibilities for types of statements about the construct. Out of a potentially limitless number of statements, we will be able to include only a fraction. The role of content validation is to provide an empirical assessment of the extent to which the items are a well-balanced and comprehensive sample of the content domain(s) in question.

How can we know that our scale items constitute a well-balanced and comprehensive sample of our content domain(s)? Alternatively and more importantly: if they do not, then

how can we identify shortcomings and amend them appropriately? These questions may be addressed by conducting a content validation study of our item pool as part of the development process. Such a study draws on a sample of experts who are knowledgeable in the content area to review the draft item pool and comment on its usefulness. It serves three related purposes:

(a) Confirmation of concept definitions

When developing multiple scales to measure multiple constructs, it is useful if expert respondents indicate how relevant they believe each item is to its associated construct. The process is akin to hypothesis testing, where a developer's thoughts about what each item measures is a hypothesis and expert responses provide (dis) confirming data. The process involves supplying respondents with a definition of each construct, and then asking them to rate the relevance of each item to the construct as defined.

(b) Clarification of ambiguous items

Experts may also comment on the clarity and conciseness of each item. This bears on reliability because even if an item is relevant, ambiguity or lack of clarity makes it less efficient at measuring the latent variable. Respondents may highlight awkward or confusing items, and suggest alternative wordings with greater clarity.

(c) Identification of additional items or areas

Respondents may be able to suggest additional items that tap the constructs well, and may also point out ways of tapping the phenomenon that the scale developer has completely missed.

Given its potential to improve the validity of the final scales, a content validation study of the attitudinal item pool was included as part of the development process.

5.5.2 Methodology

The internal validity of the CGCQ item pool was assessed by means of a postal content validation study. This section describes in detail the methodology used, substantive results, and the changes made to the draft CGCQ item pool as a result of the study. It concludes that content validation is a useful test of the internal validity of a measurement tool under construction.

The expert panel

Content validation studies rely on a nominal expert panel, charged with reviewing each item in an item pool for appropriateness and clarity. It is important therefore to identify an expert group able to make a valuable assessment of the items. The ideal composition of such an expert panel is not always obvious. A decision on the appropriateness of potential groups is required, taking into consideration their ability to contribute as well as pragmatic issues such as access to the group and their willingness to take part. In the present setting, potential expert groups included medical and/or nursing directors with a lead responsibility

for clinical governance, policy researchers with a specific interest in clinical governance, and managers with an operational responsibility for clinical governance in health care trusts. While policy researchers could provide a valuable contribution and were easily identifiable through professional networks, there were significant difficulties, relating principally to the dissemination of research information at an early stage and before peer review. Executive nursing and medical directors were easily identifiable, but there were concerns over the possibility of a low response rate from this group, and the possibility that they would pass the questionnaire to a subordinate lacking the required expert knowledge. Given these concerns, the expert group selected comprised the manager with operational responsibility for clinical governance in each hospital, community, mental health and ambulance trust in the West Midlands region. This group has a vital role in translating policy into routine practice, and routinely deals with clinical governance implementation. The decision to limit the sample to the West Midlands region was partly opportunistic, given the high profile of a clinical governance research team based in the region of which I was a member. The good relationships made with staff in the region as a by-product of this work increased the chances of a high response rate.

To ensure that the sample frame was properly drawn, I identified all eligible NHS organisations from the spring 2001 edition of Binley's directory of NHS management. This was necessary given the high level of merger activity at the time. The name of the appropriate operational manager was confirmed in a brief telephone conversation with the nurse director, or the nurse director's PA, at each trust. This served a number of purposes. Firstly, it was a necessary courtesy to the organisation, to acknowledge the support of staff

in providing information. Secondly, it confirmed the name of the most appropriate recipient within each organisation. Finally, it allowed the questionnaires to be personally addressed to the target respondents, which it was hoped would facilitate speedier returns and the highest possible response rate. Following confirmation of the correct operational manager's name, a census of all 43 managers was undertaken between April and May 2001.

Data collection

The expert sample received a personally addressed questionnaire with instructions for completion; an explanatory letter and a reply paid envelope. For each of the 29 domains, the questionnaire gave a definition, and listed the items in the pool associated with the domain. For each item, the questionnaire asked respondents 'is the item a good indicator of the domain?' and requested a response on a five-point Lickert scale ranging from 'strongly agree' to 'strongly disagree'. The response spine thus measured whether respondents thought that the statements were good indicators, not whether they personally agreed with the attitudes expressed in the items. The questionnaire also sought respondents' suggestions for alternative wordings or items and any further comments. In this way, respondents were able to state their opinion on the domain definitions, identify and clarify ambiguity in the items, and suggest additional items. The questionnaire is reproduced in full in appendix 4, and aggregated results reported in appendix 5.

5.5.3 Results

Of 43 managers in the census, 31 (72%) responded to a single postal administration, which was entirely adequate for the study purpose. When constructing the items in the pool, I was careful to follow guidelines on good practice drawn from DeVellis (1991) and Oppenheim (1992) as far as possible (table 5.4.1). Despite such caution, the validation exercise proved useful in reviewing the items under the three headings of confirming definitions, identifying ambiguity and suggesting new items.

(a) Confirm definitions

Respondents indicated the relevance of each item to its associated domain, by stating whether they agreed or disagreed that the item was a good indicator of the domain. In order to examine the status of each item, respondents' scores for each item were aggregated and organised them by domain (Appendix 5). Appendix tables list original items, respondents' scores, and changes made to items for the revised item pool.

Respondents confirmed the relevance of most of the items to the associated domains. Of 189 items, only 34 (18%) were scored as 'disagree' or 'disagree strongly' that they were a good indicator of the domain by six or more respondents. The figure of six was chosen as an arbitrary cut-off point, representing 20% of respondents. A close examination of the 34 items identified as cause for concern revealed three specific types of concern: contentiousness, ambiguity and weakness. A small number of miscellaneous items were also identified (Table 5.5.3.1).

Table 5.5.3.1: Typology of items identified by respondents as cause for concern

Cause of concern	Example original item	Re-worded item
<i>Contentiousness</i> (n)= 12	Error reporting systems are basically a stick to beat clinicians with	[No change]
	Clinicians are made to feel guilty about time away from direct patient contact	[No change]
<i>Ambiguity</i> (n)= 9	Often, improvements made elsewhere make our work more difficult	Often, service improvements made elsewhere in the organisation make our work more difficult
<i>Weakness</i> (n) = 8	We continually try to improve the quality of our practice	[Item dropped]
<i>Misc.</i> (n)= 5	Each area handles risk in its own way	[Item dropped- it was too close to another item and was thus a 'bloated particular']

The largest single category of items causing concern comprised contentiously worded items. These items sought to provoke a strong response, in order to cover the range of opinion in a domain from one extreme to another. In their detailed comments, many respondents who disagreed that such items were good indicators stated that they did not seem balanced or fair. However, the selection of contentiously worded items capable of discriminating between respondents is good practice in item construction (Oppenheim 1992). The tension for a scale designer is between avoiding loaded words, while simultaneously avoiding stilted, 'rational' statements. The decision whether to include or exclude these items was taken individually, based on a careful consideration of their potential for eliciting responses at both ends of the scale.

(b) *Identify ambiguous items*

Despite strenuous efforts during initial item construction, study respondents' identified a range of ambiguities in the draft items. These were not limited to the 34 items identified as causes of concern, as often only a very few respondents spotted a particular ambiguity. Table 5.5.3.2 outlines a typology of ambiguities contained in the draft items, with (an) explanatory example(s) and re-worded item(s). A further general problem in this area concerned the use of multiple subject descriptors in the original items, including 'I', 'we', 'people' and 'colleagues'. A number of respondents found this confusing, so the decision was taken to exclude first-person singular statements completely, and simply use 'people' or 'colleagues' wherever possible.

Table 5.5.3.2: Typology of ambiguities in the original item pool

Area of ambiguity	Example original item	Re-worded item
<i>Lack of clarity</i>	We always check the effects on others of any quality improvements we make [problem: who are the others?]	We consider the potential impact of our quality improvement activity on colleagues elsewhere in the organisation
	We have a 'not invented here' attitude to new ideas [Problem: vague]	[Item dropped]
<i>Inconsistency in wording</i>	Only a fool would admit to making an error [Problem: other items used the wording 'mistake' not 'error']	Only a fool would admit to making a mistake
<i>Wordiness</i>	Doctors <i>unilaterally</i> decide how services are run	Doctors generally decide how services are run
<i>Leading statement</i>	We continually try to improve the quality of our practice [problem: invites a 'yes']	[Item dropped]
<i>Inability to discriminate</i>	I can't see the benefit of making changes to services [problem: too obvious]	[Item dropped]
<i>Use of first-person descriptors</i>	If I made a mistake I'd report it	[Item dropped: confounds with personality type]
	I feel involved in quality improvement- we are changing rather than being changed	People are involved in quality improvement- we are changing rather than being changed

(c) Suggest new items

Despite specifically requesting suggestions for additional items, respondents suggested few. Of these, only four were of an acceptable quality for inclusion, identified below in Table 5.5.3.3

Table 5.5.3.3: New items suggested by respondents included in item pool

Domain	Item
Quality improvement: Use of cross-functional teams to address quality improvement	Staff from different clinical teams work together on quality improvement
Risk management: Assessment of risk	Clinical risk information is used routinely
Risk management: Action on identified risks	Risk management is largely a paper exercise
Appraisal: Training and development needs	Appraisals do not identify the real needs of staff

Following update of the items to address respondents' causes of concern, a number of items were re-worded to ensure roughly equivalent numbers of positively and negatively worded items in each domain. Of 189 items included in the draft item pool, 32 (17%) were excluded due to the validation study. Of the remaining 157, 65 (41%) were unchanged, and 92 (59%) had received minor changes in wording to avoid ambiguity, first-person singular statements, or ensure roughly equivalent numbers of positive and negative statements in each domain.

Conclusions

While representing a considerable investment of time, the validation study provided a number of benefits that greatly outweighed the effort. Of most value was reassurance from an expert group that the items were meaningful, and able to define the domains adequately.

If items are poor, it is much better to find out before conducting the development studies, making appropriate changes. Development samples require much larger numbers of respondents, and are more difficult to replicate should additional refinement be required of the items. Even if it leads to no changes in the items, a validation study conducted before the large-scale development sample provides valuable peace of mind to a developer.

While providing a level of reassurance, the study also highlighted a number of areas for improvement. Despite strenuous efforts during initial development, the scales included small numbers of weak, ambiguous or contentious items, requiring re-wording or exclusion from the revised item pool. This shows the importance of empirical testing before large-scale item-pool reduction work on new measurement scales.

Expert suggestions are helpful, but do not absolve the developer from final decisions on alteration or exclusion, as some suggestions run contrary to good practice in indicator development. This is particularly so in the case of the need for some redundancy in scale items, and strongly worded items capable of provoking a range of opinions. In the case of redundancy, developers need to ensure that each domain contains multiple items explaining similar ideas in a different way. The items must be sufficiently different from each other that they do not explain the same idea, yet sufficiently similar that they both act as indicators of the same latent variable. This is a difficult balance to maintain, and respondents may not understand the need to ensure that there is sufficient co-variation between items to produce stable factors. Strongly worded items may express contentious opinions on a topic that provokes strong agreement or disagreement. They are important in

scales, as they contribute variability in scores that discriminate between participants. However, respondents may not recognise the need for variability and favour bland statements instead, fearful of bias in strongly worded statements. Because of these factors, developers need to consider carefully the opinions of respondents while retaining the final decision. In this case, each item was considered on its merits and a final decision taken. Full information on the changes made to all items is included in appendix 5.

5.6 Item pool reduction

This section provides an overview of factor analytical approaches, with reference to the normative literature on its applications and advantages, and a general discussion of the applicability of factor analysis to the present study. The conclusion drawn is that its ability to reduce the CGCQ item pool and identify underlying factors capable of explaining the observed common variation among the variables makes it an ideal method. A description and analysis of the relative benefits of various forms of factor analysis follows and includes exploratory and confirmatory approaches, as well as principal components and principal factor analysis. A case is then made for the use of an exploratory principal factor model for the present study, given the need to explore the underlying factor structure of clinical governance climate, and build theory with application beyond the present dataset. Some general criticisms of factor analytical approaches are explored, including the existence of many theoretical factor models capable of fitting the data and the charge of 'garbage in, garbage out' (GIGO). There follows a detailed description of four stages in analysis; assessing the 'factorability' of the dataset; extraction; rotation; and computing factor scores.

5.6.1 Application of exploratory factor analysis

Definitions and uses of factor analysis

The normative literature identifies a number of defining characteristics and typical uses of factor analysis (Table 5.6.1.1). The term is a generic description for a range of statistical techniques concerned with examining the intercorrelations between large numbers of variables (Reymont and Joreskog 1993). By examining intercorrelations between variables, factor analytic approaches identify hypothetical underlying constructs, or 'factors', that explain the common variance between them (Rust and Golombok 1989). In explaining a complex correlation matrix in terms of a few simple underlying factors, it simplifies and describes a complex social reality (Child 1976). Factor analysis is an ideal method of indicator scale construction (Klein 1994), as it is possible to select items that load on a single underlying factor by administering indicator variables to a development sample.

In their application to test construction, the principal uses of factor analysis techniques are data reduction and the exploration or confirmation of underlying structure. Typically, this includes reducing the number of indicators in an item pool; identifying and describing a limited number of latent explanatory factors; assessing the internal consistency of sub-scales; and / or validating a theorised factor structure through hypothesis testing.

(i) *Item reduction*

By reducing the number of indicator variables in an indicator pool, factor analysis enables a parsimonious treatment of large datasets in multivariate analysis. Where there are large numbers of indicators, it identifies the subset of variables with the highest intercorrelations for retention. Reducing the number of items is important, as results from the analysis of a large number of separate variables are difficult to understand and describe. Unless they 'hang together', multiple indicators require separate analysis and reporting, which is potentially confusing and unwieldy. A further problem with a large number of separate indicators is that the need for separate statistical analysis greatly increases the number of statistical tests used. This increases the likelihood of chance 'significant' and 'insignificant' results. Consequently, it is important to reduce the number of items that are separately analysed in order to reduce the number of statistical tests required.

(ii) *Reducing variables to a limited number of explanatory factors*

This is related to item reduction above, yet goes further in its attempt to identify a limited number of factors that describe the complexity of the social world (Bryman and Cramer 1999). For example, the climate for clinical governance in healthcare organisations varies in many ways, indicated by the number of descriptions of 'good clinical governance practice' available in the literature. Yet, many of these practices seem related: the need for supportive teamwork, the avoidance of blame and learning from mistakes. If one were to collect data in the form of a series of indicator variables concerning these concepts, factor

Table 5.6.1.1: Factor analysis definitions

Child 1976:1	'Factor analysis is the orderly simplification ... of a number of interrelated measures. It should be reassuring for the reader to discover that factor analysis seeks to do precisely what man (sic) has been engaged in throughout history, that is to make order out of the chaos of his (sic) environment.'
Kim and Mueller 1978: 9	'Factor analysis refers to a variety of statistical techniques whose common objective is to represent a set of variables in terms of a smaller set of hypothetical variables.'
Norusis 1990: 149	'Factor analysis is a statistical technique used to identify a relatively small number of factors that can be used to represent relationships among sets of many interrelated variables.'
Pedhazur & Schmelkin 1991: 66	'To accept ... the validity of the indicators of the construct ... it is necessary to demonstrate that the reflective indicators of a construct (e.g. items, sub-tests) "hang together"; that they are homogenous. ... Of various approaches for studying the internal structure of a set of indicators, probably the most useful is some form of factor analysis ...'
Reymont & Joreskog 1993: 71	'Factor analysis is a generic term that we use to describe a number of methods designed to analyse interrelationships within a set of variables or objects [resulting in] the construction of a few hypothetical variables (or objects), called factors, that are supposed to contain the essential information in a larger set of observed variables or objects. ... [They] reduce the overall complexity of the data by taking advantage of inherent interdependencies [and so] a small number of factors will usually account for approximately the same amount of information as do the much larger set of original observations.'
Rust & Golombok 1989: 114	'The variables involved in the use of factor analysis are usually item scores or sub-test scores. The analysis will provide a set of results, which give an indication of the underlying relationships between the items or sub-tests. It will tell us which set of items or sub-tests go together, and which stand apart. Factor analysis identifies what are called 'factors' in the data. These factors are underlying hypothetical constructs, which often can be used to explain the data... By selecting items which relate to particular factors we are able to put together sub-tests of the construct that the factor represents.'
Tabachnick & Fidell 2001: 582	'Principal components analysis (PCA) and factor analysis (FA) are statistical techniques applied to a single set of variables when the researcher is interested in discovering which variables in the set form coherent subsets that are relatively independent of each other. Variables that are correlated with each other but largely independent of other subsets of variables are combined into factors. Factors are thought to reflect underlying processes that have created the correlations among variables.'
Tinsley & Tinsley 1987: 414	'Factor analysis is an analytic technique that permits the reduction of a large number of interrelated variables to a smaller number of latent or hidden dimensions. The goal of factor analysis is to achieve parsimony by using the smallest number of explanatory concepts to explain the maximum amount of common variance in a correlation matrix.'

analysis groups the data into a smaller number of factors. The benefit of factor analysis is that it brings a descriptive order to the way that we see things, simplifying the complex social world into a small number of hypothesised underlying constructs.

(iii) *Internal consistency*

Factor analysis identifies those variables that are highly correlated with each other. This is particularly important when researchers are interested in constructing measurement scales. Typically, measurement scales consist of a number of items, which are added together to produce a score. For the total score to be meaningful, it is essential that the items tap the same dimension. Factor analysis shows those items that tap an underlying dimension, enabling the researcher to drop proposed scale items which cross-load on multiple factors. Following the procedure, we may be confident that the overall scale score is internally consistent. Demonstration of internal consistency is particularly important when developing measures for which there is no 'gold standard' measure against which to test criterion validity (Oppenheim 1991). The assumption is that indicators in a single scale should all correlate well if they are all tapping the same latent factor.

(iv) *Assessing factorial validity*

Related to internal consistency above, factorial validity concerns the extent to which items chosen for analysis based on ability to tap a theorised latent factor actually do tap a single latent factor. If people respond in similar ways to all of the items making up a scale, then the scale has good factorial validity. The difference between this usage and internal consistency above is the *a priori* identification of theoretical latent factors. This is used in

confirmatory factor analysis, in which a hypothesised number of factors based on theoretical considerations are tested against empirical data. An analysis that confirms the hypothesised number of factors from a pool of indicator variables, and that the variables load on anticipated factors, is a stringent test of factorial validity. This approach is used where there is a well-developed conceptual framework, and a study is designed to operationalise the framework empirically.

Use of factor analysis in the present study

The data reduction and inductive theory generating properties of factor analysis identified above made its selection for use in the present study relatively straightforward. The indicator item pool of the CGCQ, developed through literature review, focus group and content validation phases described in previous chapters, needed to be reduced from over 150 items to a size that could be easily administered in healthcare settings. Further, in the absence of a well-developed theoretical framework for clinical governance climate, an exploratory factor analysis of the indicator item pool would yield hypothesised underlying factors capable of explaining the observed common variation among the identified variables. The ability of the approach to generate theory from empirical data, building on the prior qualitative work, was most valuable.

It would have been possible to select indicators in the item pool that correlated well with each other by simply examining a correlation matrix and selecting those items with high intercorrelations. However, this would have proved very difficult given the large number of variables in the pool and the resultant large correlation matrix. Further, such an

approach would not have been able to shed any light on hypothesised latent factors that could explain the correlations between the variables. For both of these reasons, a factor analytic approach was more useful in the present study context.

Exploratory Vs confirmatory approaches

Variants of factor analysis may be used either for exploratory analysis of datasets, or for more formal hypothesis testing to confirm expected underlying constructs. Exploratory factor analysis (EFA) is an inductive method of discovering underlying (latent) constructs, while confirmatory factor analysis (CFA) attempts to confirm a pre-determined structure. EFA is thus suitable for theory development, CFA for theory evaluation.

(i) Exploratory factor analysis (EFA)

Exploratory applications seek to uncover the underlying factor structure of a large set of variables. This is useful when there is insufficient empirical evidence, or lack of a clear theoretical framework, to form a hypothesis about the number of factors necessary to describe the data. There is no prior theory; rather the resultant factor loadings reveal the latent factor structure of the data. As there is no predicted factor structure, it follows that there are no *a priori* assumptions that any particular variable will be associated with any particular factor. Consequently, exploratory factor analysis is a theory generating, rather than a theory testing, procedure (Stevens 1996). EFA approaches hypothesise a limited number of factors, and identify variables that 'hang together' with others.

(ii) *Confirmatory factor analysis (CFA)*

Confirmatory applications seek to determine if the number of factors and the loadings of indicator variables on them are consistent with what is expected on the basis of pre-established theory (Kim and Mueller 1978). The rationale for indicator selection is their expected loading on a particular factor, under the assumption that each factor will load on a specified subset of indicator variables. In CFA, the test is whether indicators created to represent a hypothesised latent variable really do measure the same underlying construct.

The present study is clearly exploratory in nature. Indeed, one of the primary aims of the analysis is to explore the factor structure of the CGCQ item pool.

Variance and communality

(i) *Common and unique variance*

All factor analysis techniques are concerned with describing the variation (variance) between multiple variables in a dataset, yet they differ slightly in the manner in which they handle variance. The variance that variables share is *common variance*. All other variance is *unique variance*, of which there are two sorts. The first, termed *specific variance*, relates to that variance which is unique to a variable and not shared with any other. Finally, *error variance* is variation due to measurement error. As factor analysis is unable to distinguish specific and error variance, it combines them to form *unique variance*. Variance in variable scores is thus a compound of common, specific and error variance, with the latter two handled as unique variance:

Box 5.6.1.1: Total variance: equation 1

Total variance = Common variance + Specific variance + Error variance

When large numbers of variables are included in the analysis, common variance may spread among several factors. Where variables have common variances across a number of factors, each estimate of common variance contributes a proportion of the total variance:

Box 5.6.1.2: Total variance: equation 2 (after Child 1976)

Total variance = $C_1 + C_2 + C_3 + \dots + C_n + S + E$

Where:

C_1 - C_n = proportion of common factor variance [where n = the number of common factors]

S = specific variance

E = error variance

Each variable has an equation of this kind. A variable's *factor loadings* are simply the proportions of a variable's common variance accounted for by each factor, calculated by taking the square roots of the common variances attributable to each factor. Hence, $\sqrt{C_1}$ would be the loading of the variable on factor 1, and similarly for the other variances in the remaining factors. The basic assumption of factor analysis is that observed covariation is due to underlying common factors.

(ii) *Communality*

The sum of all the common factor variance of a variable is its *communality* (h^2), the variance that a variable shares in common with other variables. As each factor loading is simply the square root of the common variance attributable to each factor, a variable's communality is also the *Sum of the Squares of its common factor Loadings* (SSL).

Ideally, the common variance of each variable should be large and accounted for by one factor. By examining intercorrelations between variables, factor analysis identifies those variables that load highly on only one factor. Items retained for scales require high communalities (shared variance) and a high factor loading on a single factor (i.e. a high proportion of shared variance attributable to a single factor). For this reason, it is conventional to report communalities and factor loadings for all item variables in factor analysis results.

Factor Analysis (FA) Vs Principal Components Analysis (PCA)

The two basic models for factor solutions are factor analysis (FA) and principal components analysis (PCA). In practice, the major distinction between them lies in the way that they handle unique variance (Bryman and Cramer 1999). FA takes account of the presence of unique variance and attempts to exclude it, using only common variance in solutions. In contrast, PCA assumes that all variance is common, and produces a solution based on all of the variance, including unique variance. They handle unique variance so differently because of a difference in the nature of the assumed association between variables and factors on the one hand, and variables and components on the other

(Tabachnick and Fidell 2001). In FA, the assumption is that underlying factors 'cause' variables; that is, the presence or absence of underlying factors causes scores on variables. Thus in exploratory FA the question is what underlying causative processes produce the correlations between the variables. In contrast, components are simply aggregates of correlated variables; the variables 'cause' the components. In PCA, variables and components are simply empirically associated; there is no theory about the underlying nature of the associations. Components are simply a shorthand way of describing empirical correlations within a single dataset in order to simplify multivariate analyses; they have no validity beyond a particular dataset.

Choice of exploratory factor analysis (EFA) model in the present study

From the preceding discussion, factor studies may be characterised by their position on two axes: exploratory or confirmatory, and components or factors. The aim of the study is to explore the underlying factor structure of indicator variables designed to reflect the climate of clinical governance in health care trusts. This suggests a theory-building, exploratory approach to the underlying latent variables. There is no prior hypothesis; indicator variables were included for their importance to the universe of interest, but not expected to load on a pre-defined latent factor. The intention to build theory beyond the present dataset led to the decision to examine only common variance between the variables, necessitating a factor analysis approach. Consequently, the study utilised an exploratory factor analysis (EFA) model.

Factor analysis: royal road to truth or shotgun empiricism?

The proper use of EFA is concerned with the question of the smallest number of factors required to adequately summarise the relationships among a given set of indicators (Pedhazur and Schmelkin 1991). Yet, exploratory approaches have attracted vehement criticism. The main objective of factor analysis, the identification of an underlying factor structure, is not straightforward. This is because there are a potentially limitless number of causal structures (factor models) that could produce the same pattern of variation between variables. The basic problem is one of indeterminacy between correlation matrix and factor structure. Different causal structures that could share a correlation matrix include non-factor models with no underlying causative factors, factor models with different numbers of common factors, and factor models with different factor loadings (Kim and Mueller 1978). This apparent ability of factor analysis to fit almost any solution led to a widespread general dissatisfaction with the technique in the 1970s (Rust and Golombok 1989). However, there are three basic assumptions used to minimise uncertainty over the appropriateness of a given factor solution, and together they offer a powerful corrective to criticism of factor analysis as 'shotgun empiricism': factorial causation; parsimony; and rotation.

Factorial causation

This first assumption postulates that observed variables are linear combinations of underlying factors, rather than simply a melange of data, and is used to refute the charge that there is no underlying factor structure. The results of a factor analysis cannot be used

to make the case; this would be simple circularity. Results simply show that the structure of the data is consistent with a factor model based on such a postulate (Kim and Mueller 1978). The charge is that EFA is essentially a blind process, in which a large number of disparate variables are placed into an analytic grinder in the hope of generating 'interesting' results, known as 'garbage in, garbage out' (GIGO). While open to such abuse, properly conducted EFA is able to rebuff the charge. A well-designed EFA includes indicators chosen from the universe of interest because of their importance, rather than simply including variables at random. In the present study, indicators describe the lived experience of professionals in relation to clinical governance. They are derived from substantial qualitative development and further validation with an expert panel, and there is a strong case for their inclusion in the analysis. This development process drew on a review of elements of clinical governance, so that one might reasonably expect the presence of underlying latent factors.

Parsimony, or 'simple solution'

This principle addresses the problem of models with different numbers of factors that are consistent with the data. It favours the simplest solution i.e. the model with the smallest number of factors that is consistent with the observed data. The practical import of the approach is that it provides a way of selecting a 'favoured' solution. While there are many potential factor solutions to a correlation matrix, there is only one configuration consistent with the postulate of the smallest number of factors that fits the data.

Rotation

Following identification of the number of factors, rotation addresses the final indeterminacy of factor loadings. Rotated solutions all explain as much covariation in the data as the original solution and each other. They are attempts to simplify the solution, making it more interpretable. Many different models are available, but they typically clarify which variables are associated with a given factor, or which factors are associated with a given variable. Options are discussed below, in the context of the present study.

Operationalising a parsimonious factor solution and rotating to final solution

A successful factor solution is one that simplifies the correlation matrix and produces meaningful factors. That is, one that explains the observed correlation between variables using as few factors as possible, and provides meaningful factors that are easily interpretable. There are usually four steps in factor analysis (Norusis 1990):

- (i) Assessment of the 'factorability' of the data. This involves examination of the correlation matrix, to identify and exclude variables that are unrelated to each other. In addition, the Kaiser -Meyer-Olkin measure of sampling adequacy (KMO MSA) measure of sampling adequacy and Bartlett's test of sphericity may further assess the appropriateness of the matrix for factor analysis.
- (ii) Factor extraction. At this stage, the method of calculating the number of factors necessary to represent the data is determined, and calculations made. Also required is an assessment of the degree of 'fit' between the chosen model and data.
- (iii) Factor rotation. This transforms the initial factor solutions in order to increase their interpretability.
- (iv) Computation of factor scores for each case, used in further analyses.

(i) *Assessment of the 'factorability' of the data*

In order for the factor model to be appropriate, the variables in the analysis must correlate well with each other; if the correlations between variables are low, it is unlikely that they share common factors. Examination of the correlation matrix reveals whether there are large shared correlations (>0.3) between variables (Norusis 1990). The statistical test for whether there are no significant correlations between variables is Bartlett's (1954) test of sphericity, which requires a large test statistic and a small associated significance level (<0.05). This test statistic requires normally distributed data and is notoriously sensitive to sample size. It is so sensitive that with large sample sizes it is likely to be significant even if the correlations are low, so is recommended only when there are fewer than five cases per variable (Tabachnick and Fidell 2001).

The KMO MSA provides an additional test of factorability. If variables share common factors, then the partial correlations between pairs of variables with the linear effects of other variables removed should be low. The KMO MSA is a ratio of the sum of squared correlations to the sum of squared correlations plus partial correlations, and its value approaches 1 when partial correlations are small. Values of 0.90 are marvellous, 0.80 meritorious, 0.70 middling, 0.60 mediocre, 0.50 miserable and below 0.5 unacceptable (Kaiser 1974). SPSS reports both a global KMO MSA statistic, and a KMO MSA for each variable in the diagonal of the anti-image correlation matrix. Variables with low KMO MSA are redundant as they do not share common factors with other variables, and Norusis (1990) recommends their removal from the analysis.

Given the recorded sensitivity of Bartlett's test of sphericity, this study additionally reports the KMO MSA for the sample and individual variables. The results section identifies and lists redundant variables.

(ii) *Factor extraction*

The goal of factor extraction is to determine the optimum number of factors. PCA extracts linear combinations of observed variables in sequence, starting with the combination accounting for the largest amount of variance. Successive components account for progressively smaller amounts of the total variance, and each component is uncorrelated with every other.

The optimum number of components to extract is a trade-off between proportion of common variance accounted for and simplicity of solution. They are in tension because the more components extracted the larger the proportion of common variance explained by the solution; yet, the more complex the solution. The two principal methods for assessing the optimum number of factors are Kaiser's criterion and Cattell's scree test.

Kaiser's criterion

This method relates to the percentage of total variance explained by each factor. PCA expresses variables and components in standardised form, with a mean of zero and standard deviation one. As each variable has a standardised variance of one, so the total variance reported is equal to the number of variables in the analysis. The amount of variance

explained by each component is its *eigenvalue*, also expressed in standardised form. Thus, a factor with an eigenvalue of two accounts for twice as much variance as a single variable. The percentage of variance accounted for by an eigenvalue is given as:

$$100 / \text{Number of variables} * \text{eigenvalue}$$

Thus, in a solution with 50 variables, a factor with an eigenvalue of 22 accounts for 44% of the variance.

Because the variance contributed by each standardised variable is one, any component with an eigenvalue of less than one explains less variance than a single variable. Kaiser's criterion suggests retention of all components with eigenvalues >one. However, there are difficulties associated with this approach, especially concerning its use with large numbers of indicator variables. Stevens (1996) recommends Kaiser's criterion when there are less than 30 variables and the average communality is above 0.70, or when there are over 250 participants and the average communality is above 0.60. Child (1976) suggests that with less than 20 variables the Kaiser criterion extracts too few factors, and above 50 variables it extracts too many. Under these circumstances, Child (1976) recommends Cattell's scree test as an alternative.

Cattell's scree test

Cattell (1966) argues that the proportion of unique variance in later factors increases, to the point of swamping common variance. Thus, the extraction of too many factors is doubly serious; not only does it lead to an over-complex solution, but later factors are increasingly unreliable, and so may not be replicable with other samples. For these reasons, it is important to identify an optimum number of factors before the intrusion of unique variance becomes serious.

Cattell's scree test produces a more parsimonious solution than Kaiser's criterion, suggesting fewer factors. The test plots eigenvalues against factor number, with factor numbers arranged in decreasing order. The plot is negatively decreasing, with eigenvalue highest for the first factor and moderate but decreasing for the next few factors, before tailing off for the final factors. To select the number of factors, one looks for the point at which a line drawn through the points noticeably changes slope.

The scree test is not an exact test; it involves a judgement of where the discontinuity in eigenvalues begins. Tabachnick and Fidell (2001: 622) recommend an initial analysis and scree plot to ascertain an approximate number of factors for retention, then further analyses, each specifying a different number of factors either side of the initial estimate. Examination of the residual correlation matrix for each of the subsequent analyses shows the partial correlations between variables with the effects of factors removed. If the analysis is good, the residuals should be small. If there are a few large residuals ($>.10$) or many moderate residuals ($>.05$), then a further unextracted factor is suggested. The model

with the smallest number of factors that satisfies the residual analysis is reported, a balance between simplicity and explanatory power. The results section below details the number of factors identified by Kaiser's criterion on an initial analysis, a scree test plot, and residual analysis of a number of subsequent analyses before delivering a final decision on factor extraction.

Significance of factor loadings

The higher the loading, the more the variable is a pure measure of the underlying factor. Comrey and Lee (1992) suggest loadings more than .71 are excellent, .63 very good, .55 good, .45 fair and .32 poor. In practice, the choice of cut-off is largely arbitrary, and if there is a gap in loadings consistent across factors, it is the default cut-off point.

(iii) Factor rotation

Although the extraction factor matrix indicates the relationship between the factors and individual variables, it is usually difficult to identify meaningful factors. As this is usually one of the primary goals of analysis, *rotation* transforms the initial matrix into one that is easier to interpret. It is important to stress that rotation improves only the interpretability of the solution; it does *not* improve the 'fit' between observed and reproduced correlation matrices. Residual analysis at the factor extraction stage (cf. above) assesses the adequacy of the factor solution to account for empirical variation in the dataset.

Klein (1994) notes the existence of a 'general factor' in FA and PCA that is an artefact of the algebraic procedures. He suggests that the best way round this in exploratory FA is to

develop several scales at once, and advises developers' to include variables believed to relate to a number of different discrete factors in a single analysis. As rotation techniques redistribute variance from the general factor to subsequent factors, rotation to simple structure makes sense in these circumstances, where variables relate to one of several scales, each factor representing a different scale.

The two generic approaches to rotation are *orthogonal* and *oblique* techniques, and they strike a different balance between conceptual sophistication and ease of interpretation. Orthogonal rotations produce uncorrelated factors that are relatively easy to interpret and describe. The difficulty is that unless one is sure that the underlying processes are unrelated they may stretch credibility. While there are conceptual advantages to oblique rotations where there is reason to believe that the underlying factors are correlated, these are offset by increased practical difficulties in interpreting results.

Kim and Mueller (1978) argue that in exploratory factor analysis, where the emphasis is on identifying the basic structuring of variables, the selection of rotation is a secondary consideration, and any readily available method will do. They suggest that correlation between factors may be of little import, and counsel for orthogonal rotations on grounds of pragmatism i.e. that orthogonal rotation is easier to interpret. Klein (1994) agrees, stating that orthogonal solutions revealing an interpretable solution are preferable to oblique alternatives given their greater simplicity. However, Cattell (1978) and Thurstone (1947) are critical of pragmatic arguments for orthogonal rotation, noting the frequency of correlations between underlying constructs. Child (1976) sums up this perspective well:

'Oblique rotation with behavioural variables is an admission that most, if not all, human characteristics are correlated to some extent, and the underlying major factors must be similarly correlated.' (Child 1976:60)

From this perspective, a decision in favour of orthogonal rotation, however understandable, may not be justified, as it misrepresents the real world. This may be missing the point however, as the defence of orthogonal techniques is made on grounds of their simplicity and ease of replication across studies, rather than their ability to represent an external 'reality'. Thus in circumstances where the aim of the study is to identify stable factors, orthogonal rotation may be both simpler and more reliable.

The similarity between orthogonal and oblique solutions where patterns of correlations in the dataset are clear (Tabachnick and Fidell 2001) makes orthogonal rotation a good choice. Pedhazur and Schmelkin (1991) counsel for both rotations, and if correlations among factors is small, reporting of the orthogonal solution. Tabachnick and Fidell (2001) agree, specifying correlations of above 0.32 (10% of the variance) as significant. However, Rust and Golombok (1989) urge caution in the use of oblique rotation techniques, and given that variations in the extent to which orthogonal requirements are relaxed produce often dramatically different factor solutions they warn against its use. Both oblique and orthogonal rotations are considered in the results section 5.6.3 below.

Many rotation algorithms exist, designed for different ends. In exploratory factor analysis, the most common orthogonal rotation procedure is varimax rotation, which simplifies factors by maximising the spread in factor loadings and redistributing variance from earlier

to later factors (Tabachnick and Fidell 2001). This makes interpretation of factors easier because it is more obvious which variables are associated with which factors. As the identification of latent factors is a primary goal of the present analysis, it is the most appropriate orthogonal method. An oblique alternative is promax rotation, which takes a varimax solution and rotates it a second time to allow correlations between the factors. The advantage of the promax solution over other oblique alternatives is that simple structure is maximised, as the initial varimax stage clarifies which variables correlate with which factors.

(iv) *Computation of factor scores*

A number of options are available for factor score compilation. The simplest method simply adds together raw item scores of each item loading on a factor. This has the advantages of transparency and simplicity, and allows subsequent users of a questionnaire schedule to compute factor scores for comparative purposes. While such an approach has obvious appeal in the present context, there are other options. Regression factor scores, computed as part of the SPSS FACTOR procedure, produces uncorrelated normally distributed estimates. This has advantages when factor scores are used as dependent variables in subsequent multivariate analysis, given the assumption underlying multiple analysis of variance that variables are normally distributed, and difficulties in overestimating effects of independent variables when dependent variables are correlated. Consequently, simple summated factor scores are reported for CGCQ domains in section 5.7 on reliability and validity. Given their statistical advantages, uncorrelated regression factor scores are used in the multiple analysis of variance (MANOVA) to explore the

effects of professional group and organisation on perceptions of clinical governance climate reported in section 6.2.3.

5.6.2 Methodology

This section describes the methods used to reduce CGCQ indicator item pool, and provides results of an exploratory factor analysis. Results of additional reliability and validity testing are reported in section 5.7 below.

Aim

The aim was to reduce a pool of clinical governance climate indicators developed via prior qualitative research and describe a latent factor structure of clinical governance climate for the CGCQ. The study additionally reports the internal consistency and external validity of the factor model in a defined population.

Methods

Sample size

As correlation coefficients are less reliable when estimated from small samples, it is important that sample size is large enough to estimate correlation coefficients reliably. There is no consensus on what the size of the sample should be (Bryman and Cramer 1999). As a rough guide, Comrey and Lee (1992) identify sample sizes of 100 as poor, 200 fair, 300 good, 500 as very good, and 1000 excellent. As solutions with relatively low

loadings need larger sample sizes (Guadagnoli and Velicer 1988), caution dictates as large a sample as is realistically obtainable.

Samples were drawn from four health care trusts: an acute teaching hospital; a Primary Care Trust (PCT); a community, mental health and learning disabilities trust; and a specialist acute hospital. Trusts were offered free, anonymised comparative information of scores on the CGCQ domains, which was considered valuable. Each of the four trust sites drew a random sample of 210 staff from HR records, distributing questionnaires through internal mail systems. In order to boost sample size, an additional sample of 300 healthcare managers was drawn from recent conference attendees at the Health Services Management Centre, University of Birmingham, who received questionnaires mailed to their place of work. A total of 1140 questionnaires were mailed. Respondents returned completed questionnaires directly to the researchers. In the present study, data are available from 406 subjects with no missing values for indicator items.

Procedures

Conducted between December 2001 and May 2002, the study involved staff at four healthcare trusts and an additional sample of healthcare managers employed in the UK healthcare sector. Sampled participants received a single mailing of a ten-page questionnaire containing indicators of clinical governance climate, demographic and other measures, accompanied by a covering letter explaining the study purpose and assuring respondent confidentiality.

Study measures

As detailed above, the indicators required to build a factor model of clinical governance climate were developed qualitatively through a series of nine focus groups of health service professionals, involving 65 staff in four health care trusts. Participants included senior medical, nursing and managerial staff, who identified ideal and current practice in their organisations in each of the nine domains identified earlier. Content analysis of the discussions generated over 160 potential indicator items of clinical governance climate related to the domains identified in the document review. Items consisted of positive and negative statements in participants' own words, designed to be scored as Lickert-scales in which respondents identify the strength of their agreement or disagreement with each statement. Following content validation, the draft item pool of the Clinical Governance Climate Questionnaire (CGCQ) contained 157 indicator items.

Statistical methods

Data were analysed using SPSS-PC for Windows software (SPSS Software Inc, release 9.0). Exploratory principal factor analysis (EFA) was used to empirically derive the separate dimensions underlying the CGCQ item pool and identify items for retention based on shared variance (Streiner and Norman 1995). As factor analysis uses only the common variance between variables, it is particularly appropriate in applications seeking to build theory inductively with application beyond a single dataset (Tabachnick and Fidell 2001). Internal consistency was further assessed using Cronbach's alpha (Cronbach 1951). The strength and direction of correlations between sub-scale scores for construct validation were assessed using Pearson's correlation (r). Differences in mean factor scores between

professional groups and trusts were examined using one-way (single factor, non-repeated measure) analysis of variance, reporting F-tests for significance.

Data cleaning and preparation for factor analysis

To improve the quality of analysis, cleaning procedures were adopted and the assumptions of normality, linearity and homoscedacity underlying multivariate analysis tested (Table 5.6.2.1). Checking for out of range values and plausible distributions supplemented by further examination of 40 (10%) entered cases against the original paper questionnaires was used to assess data entry accuracy. Examination of the extent and distribution of missing values and estimation of replacements in order to retain cases followed. Assessment of the extent of departure from normal distribution, linearity and homoscedacity concluded with appropriate transformations to approximate normality, and the exclusion of twenty-six poorly discriminating variables. Screening revealed no multivariate outliers with undue influence on results, and no redundant variables with perfect (singular) or near-perfect (multicolliniar) correlations.

Table 5.6.2.1: Data screening (after Tabachnick and Fidell 2001)

Identify	Rationale
(a) Univariate descriptive statistics: • out of range values • plausible distributions (b) Amount and distribution of missing data. (c) Univariate outliers (d) Non-normal variables (skewness and kurtosis) (e) Check multivariate assumptions: • non-linearity • heteroscedacity (f) Make appropriate mathematical transformations to better approximate normality (g) Multivariate descriptive statistics • multivariate outliers • variables causing multivariate outliers • description of multivariate outliers (h) Evaluate for multicolliniarity / singularity	Check accuracy of data entry Reduce amount of missing data in analysis; exclude cases / variables or estimate values for missing data Identify factors distorting correlations Assumption of normality Assumptions underlying multivariate analysis Meet assumptions of analysis Check for undue influence of cases on results Screen for 'perfect' relationships between variables and remove redundant ones

(a) Univariate descriptive statistics for out of range values and plausibility

Data entry accuracy for all 406 cases was assessed through the SPSS FREQUENCIES command for categorical data. All data was within range, and missing value codes properly assigned. Subsequent cross-tabulations of variables for educational qualifications and professional role showed logical results: e.g. medical practitioners had medical qualifications; nursing practitioners had nursing qualifications. Entry validation, assessed by an examination of the first 40 (10%) cases of entered data against the original paper questionnaires, showed that all (201 items x 40 cases = 8040) data points were correctly entered.

(b) Amount and distribution of missing data

Missing data refers to questions left blank by the respondent. The seriousness of the problem depends on the proportion of missing data relative to the total dataset, and its pattern. If 5% or less of data points are missing from a large dataset, the result is not sensitive to the procedure for handling missing data (Tabachnick and Fidell 2001). While the amount (proportion) of missing values is important, it is not as critical as its pattern, as non-random distribution of missing values affect the generalisability of results. For example, if a number of doctors in the sample refuse to answer a question on team co-operation, estimates of team co-operation will not be generalisable to doctors.

The basic choice is between deleting cases and /or variables with missing data, or imputing values for insertion into the dataset. Random distribution among a few cases and on different variables suggests the deletion of those few cases. However, this would not be appropriate where the distribution is across many cases and / or variables, resulting in the loss of many cases and / or variables to analysis. This latter situation requires the imputation and insertion of scores on cases with missing values.

SPSS Missing Values Analysis (MVA) provides a test for proportion and patterning of missing values, computing a t-test for any variable with at least 5% of data missing against every other variable. In the current dataset, SPSS MVA identified no variables with more than 3.3% (13 cases) of data missing. As missing values were distributed across sixty-seven cases, deletion of all cases with missing data would reduce cases available for analysis from 406 to 339. As this amounted to a significant reduction, missing values were

overwritten with imputed scores using SPSS MVA Expectation Maximisation (EM), ensuring the availability of all 406 cases for analysis.

(c) Univariate outliers

Univariate outliers are cases with extreme values on a variable that distorts statistics, often caused by errors in data entry or specification of error values but also due to extreme non-normal distributions in continuous variables. As indicator variables were coded as Lickert scales between 1 and 5 and univariate descriptive statistics had shown that all were coded within range, there were no univariate outliers in the dataset. As items consisted of positively and negatively worded statements scored between 1 ('agree strongly') and 5 ('disagree strongly'), positively worded statements were recoded, using the transformation $\text{Newscore} = (6 - \text{Oldscore})$. This meant that positively and negatively worded items loading on the same factor could be added together to produce a summary score, with high scores meaning good performance, and low scores meaning poor performance.

(d) Normality in distribution

While normal distribution of variables is not critical in PCA and FA solutions where the emphasis is on a descriptive summary of relationships between observed variables, it does enhance the solution (Tabachnick and Fidell 2001:588). Normality among single variables is most often assessed by skewness (asymmetry of distribution) and kurtosis (peaking), either graphically via histograms or via test statistics. Values of zero indicate normality and variables with substantial skewness and / or kurtosis may be mathematically transformed to better approximate normal distribution.

In the present dataset, departures from normality were assessed by using the SPSS EXPLORE function to request histograms and summary statistics for skewness / kurtosis. The analysis revealed many negatively skewed and some positively skewed distributions (Appendix 8).

(e) Non-linearity and heteroscedacity

Non-linearity

Linearity assumes straight-line relationships between variables. As correlation measures linear associations, factor analysis solutions are degraded when this assumption fails.

Homoscedacity

The assumption of homoscedacity is that the variability in scores for on continuous variable is roughly the same at all values of another continuous variable. Analysis is weakened, although not fatally, to the extent that there is heteroscedacity between variables.

Because of non-normality in skewness and kurtosis values, the extent of departure from linearity and homoscedacity was examined via bivariate plots. Continuous data requires inspection of bivariate scatterplots, which are oval where variables are linearly related. As the variables in the current dataset are ordinal with few categories, bivariate barcharts were examined using SPSS CROSSTABS. Skewness and kurtosis statistics were examined to identify pairs likely to depart from linearity i.e. a highly positively and highly negatively skewed pair. Figure 5.6.2.1 shows bivariate barcharts for the moderately skewed items 17

and 151; the normally distributed items 15 and 21; and the highly positively skewed item 61 against the highly negatively skewed item 120. The asymmetry of distribution in barchart 3 is evident in the pileup of scores under the negative end of N120 and the positive end of N61. These variables are not linearly related. Heteroscedacity is also evident in the greater variability in N61 scores for negative than positive values of N120.

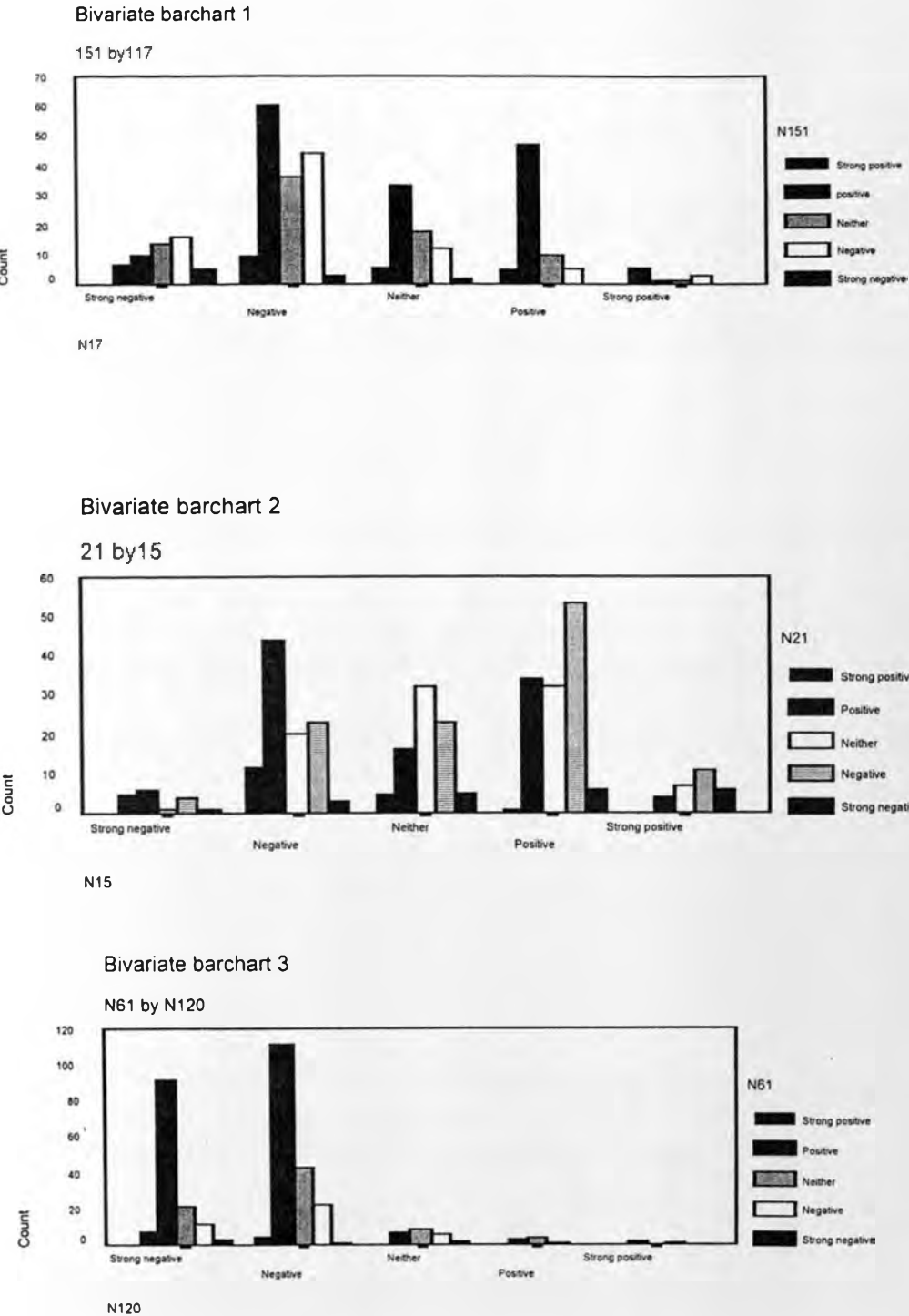
(f) Transformations

Given the above evidence of non-normality, non-linearity and heteroscedacity in the dataset, variables were transformed to better approximate normal distribution, using logarithms for highly skewed, and square roots for moderately skewed, items. Post-transformation skewness and kurtosis statistics show much improved statistical evaluations (Appendix 8). Twenty-six variables with limited variation and thus poor discrimination were excluded from further analysis, again identified in Appendix 8.

(g) Multivariate outliers

Cases with extreme outlying values on combinations of variables have more influence on factor solutions than others. Datasets may be screened for outlying cases by calculating a mahalanobis distance, which is the distance of a case from the point of intersection of the means of all variables for all other cases. Each case is represented by a point based on its combination of scores on all variables, with outlier cases lying some distance from the 'swarm' of other cases (Tabachnick and Fidell 2001:68). Transformed cases were screened for multivariate outliers using SPSS REGRESSION, specifying RESIDUALS=OUTLIERS

Figure 5.6.2.1: Bivariate barcharts of CGCQ item distributions



MAHAL), with case id number as the dependent variable. As no cases were identified as outliers, all 406 were retained for analysis.

(h) Multicollinearity and singularity

Variables that are too highly correlated (multicollinear) or are a direct linear combination of two or more of the others (singular) contain redundant information and should be dropped from analysis. Investigation involves examination of the Squared Multiple Correlations (SMCs) for each variable as it serves as a dependent variable with all others as independent variables. In factor analysis, SMCs are also known as communalities i.e. the variance that a variable shares in common with other variables. While communality needs to be moderately high to denote shared variance, values more than 0.9 may denote multicollinearity and values of 1 denote singularity. As examination of SMCs revealed no value greater than 0.8 singularity and multicollinearity are not problems in this dataset.

5.6.3 Exploratory Factor Analysis (EFA) results

Respondent characteristics

Of 1140 questionnaires, respondents returned 406 (36%) from a single mailing. While this is disappointing, it is consistent with other reported response rates from unsolicited questionnaire mailings. Table 5.6.3.1 provides descriptive information on sample respondents in terms of professional group, professional seniority, and length of time at

present trust and mean age. The sample included respondents from medical, nursing, therapy and managerial staff groups. While senior staff from each group were well represented, there were few responses from junior medical staff.

Table 5.6.3.1: Respondent characteristics

Variable	(n)	(%)
<i>Professional group</i>		
Medical	58	14.3
Nursing / therapy	151	37.2
Managerial	129	31.8
Other	65	16.0
Missing	3	0.7
<i>Years at present trust</i>		
<1 year	71	17.5
1-4 years	112	27.6
5-9 years	62	15.3
10-14 years	69	17.0
15+ years	89	21.9
Missing	3	0.7
<i>Age</i>		
21-29	62	15.3
30-39	120	29.6
40-49	145	35.7
50-59	68	16.7
60+	11	2.7
<i>Professional seniority</i>		
Consultant / specialist	44	10.8
Registrar	3	0.7
SHO / HO	1	0.3
GP	10	2.5
Nurse 'F' to 'H'	70	17.2
Nurse 'C' to 'E'	21	5.2
HCA	9	2.2
Nurse other	5	1.2
Senior therapist / psychologist	28	6.9
Therapist / psychologist	8	2.0
Therapist other	10	2.5
Executive director	7	1.7
Divisional manager / Directorate general manager	67	16.5
Deputy manger	8	2.0
Manager other	47	11.6
Admin / clerical	49	12.0
Technical / craft	8	2.0
Other	8	2.0
Missing	3	0.7

Factorability of the item pool

As the purpose of factor analysis is to explain the common variance between indicator variables, factorable matrices should contain a number of high correlations between variables. A correlation matrix of the 157 CGCQ items produced by SPSS FACTOR revealed a large number of correlations higher than 0.30, which suggests the likelihood of patterns in responses to the variables. Bartlett's test and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO MSA) were requested through SPSS FACTOR to further assess item factorability. Bartlett's test of sphericity is highly significant (Chi-sq.= 30820, df. =8385, $p < .001$), suggesting that it is extremely unlikely that the correlations in the sample correlation matrix are due to chance. Kaiser's measure of sampling adequacy (MSA=. 949) is very close to one and strongly suggestive of underlying factors. Additionally, a check of the anti-image correlation matrix revealed no variables with an individual MSA of <0.6 , again suggesting good factorability.

(i) Factor extraction

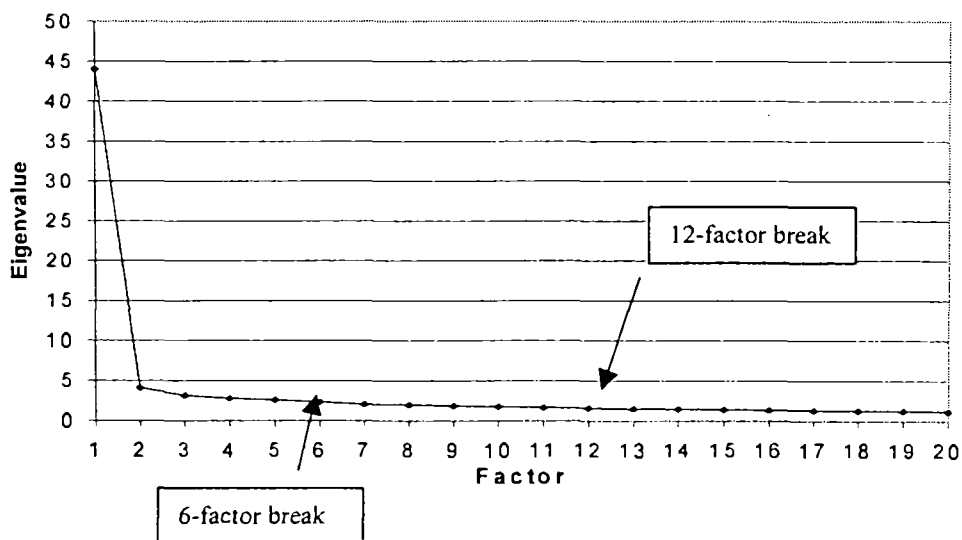
An exploratory factor analysis (EFA) extraction with varimax rotation through SPSS FACTOR was used as an initial run to estimate the number of latent factors from eigenvalues. Table 5.6.3.2 shows the first 20 eigenvalues. The maximum number of factors with eigenvalues >1 is 17. However, retention of all 17 factors seems unreasonable, so sharp breaks in the size of eigenvalues were sought using Cattell's scree test (Figure 5.6.3.1).

Table 5.6.3.2: Total Variance Explained, initial 20 factors

Factor	Initial Eigen			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	44.035	33.873	33.873	43.640	33.569	33.569	8.611	6.624	6.624
2	4.105	3.158	37.031	3.717	2.860	36.428	8.221	6.324	12.947
3	3.165	2.435	39.465	2.736	2.104	38.533	8.053	6.195	19.142
4	2.810	2.162	41.627	2.391	1.839	40.372	4.882	3.755	22.898
5	2.595	1.996	43.623	2.165	1.666	42.038	4.696	3.612	26.510
6	2.339	1.799	45.422	1.941	1.493	43.531	4.188	3.221	29.731
7	2.054	1.580	47.002	1.634	1.257	44.788	3.825	2.942	32.673
8	2.000	1.538	48.540	1.585	1.219	46.007	3.422	2.633	35.306
9	1.850	1.423	49.963	1.447	1.113	47.120	2.888	2.221	37.527
10	1.768	1.360	51.323	1.349	1.038	48.158	2.698	2.075	39.603
11	1.653	1.271	52.595	1.239	.953	49.111	2.322	1.786	41.389
12	1.581	1.216	53.811	1.186	.913	50.023	2.266	1.743	43.131
13	1.478	1.137	54.948	1.053	.810	50.833	2.137	1.644	44.775
14	1.462	1.125	56.073	1.041	.801	51.634	2.102	1.617	46.392
15	1.409	1.084	57.157	.974	.749	52.383	2.046	1.574	47.966
16	1.327	1.021	58.178	.909	.699	53.082	1.992	1.533	49.499
17	1.313	1.010	59.188	.857	.659	53.741	1.591	1.224	50.722
18	1.262	.971	60.159	.843	.649	54.390	1.450	1.115	51.838
19	1.252	.963	61.122	.817	.628	55.018	1.317	1.013	52.851
20	1.194	.918	62.040	.773	.595	55.613	1.311	1.008	53.859

Extraction Method: Principal Axis Factoring.

Figure 5.6.3.1: Scree-plot of factors against eigenvalues



Eigenvalues for the first 3 factors are all >3 and, after the 12 factor, changes in successive eigenvalues are small (table 5.6.3.1). This suggests between 3 and 12 factors. The scree plot visually suggests breaks between 6 and 12 factors. Identification of the optimal number of factors required several EFA runs specifying from 6 to 12 factors.

To remove unique and error variability from each variable and select the optimal number of factors, a common factor extraction model (Principal Factor Analysis) was used for several runs specifying between 6 and 12 factors. Specification of the correct number of factors is important, as if there are too few factors in the model then variance remains unaccounted for, whereas retention of too many factors leads to fragmentation. As the analysis was principally concerned with identifying reliable factors, the goal was to retain the smallest number of interpretable factors consistent with residual analysis. The residual correlation matrix was used to identify partial correlations between pairs of variables with the effect of factors removed, as high residuals values suggests the presence of additional unextracted factors in the dataset not contained in the model.

Table 5.6.3.3 identifies partial correlations and variance accounted for by each of the EFA models considered and additionally reports the number of factors in each model defined by more than three variables. This latter is important as factors with fewer than three variables loading on them are poorly defined, especially if they correlate with other variables (Tabachnick and Fidell 2001). Above eight factors, there is evidence of some fragmentation, with some late-extracted factors having fewer than four loading variables.

Table 5.6.3.3: Factor model adequacy

Criterion	Number of factors in the EFA model						
	6	7	8	9	10	11	12
Partial correlations >0.05 (%)	17	15	13	11	10	9	8
Variance explained (%)	42.8	44	45	46.3	47.3	48.2	49.1
No. of factors >3 variables	6	7	8	8	9	9	10

The six-factor solution accounts for over 40% of the variance and satisfies residual analysis, as most (83%) values in the residual correlation matrix are near zero. As the goal was to identify the smallest number of interpretable factors, the six-factor solution was selected for full reporting.

(ii) Rotation

The decision between oblique and orthogonal rotation was taken after requesting principal factor extraction with oblique rotation for six factors and examining the factor correlation matrix for correlations >0.32 (10% of the variance). Promax rotation with gamma=2 showed moderate correlations of around .32 between factors 1 and 2, and 2 and 3. (Table (5.6.3.4). While this level of correlation between factors is borderline for an orthogonal solution (Tabachnick and Fidell 2001), there are considerable advantages to orthogonal factors. Crucially, orthogonal factor scores may be used in further multivariate analysis without problems of confounding between scales. Consequently, the final model reports the simpler orthogonal solution. The model evaluated, interpreted and reported is thus the six-factor model with principal factors extraction and Varimax orthogonal rotation.

Table 5.6.3.4: Factor Correlation Matrix

Factor	1	2	3	4	5	6
1	1.000					
2	.348	1.000				
3	.310	.321	1.000			
4	.300	.242	.203	1.000		
5	.265	.268	.177	.208	1.000	
6	.307	.257	.220	.195	.189	1.000

Extraction Method: Principal Axis Factoring.

Rotation Method: Promax with Kaiser Normalization.

(iii) *Model adequacy*

The adequacy of EFA models may be assessed both statistically and with regard to the interpretability of emergent factors.

Variables well defined by the solution?

Communalities, which indicate the percentage of variance in a variable overlapping variance in the factors, show whether the solution well defines the variables (Table 5.6.3.5).

Twelve are lower than 0.3, indicative of some heterogeneity among the variables.

However, the variables are reasonably well defined by the solution.

Importance of each factor

The importance of each factor is assessed by the percentage of variance and covariance it accounts for. The former is calculated by dividing the variance accounted for by the factor (Rotated Sum of Squared Loading SSL) by the number of variables, and multiplying by 100. The latter is calculated by dividing SSL by the sum of SSLs and multiplying by 100

(Tabachnick and Fidell 2001). The factors account for between 10.6 and 5.5% of the variance (Table 5.6.3.5).

Table 5.6.3.5: percentage of variance and covariance explained by rotated factors

	Factors					
	1	2	3	4	5	6
Rotated SSL	13.8	10.6	9.2	7.5	7.5	7.1
% variance	10.6	8.1	7.0	5.8	5.7	5.5
% covariance	24	19	16	14	14	13

Total variance explained by six-factor model = 42.8%

Simplicity of structure

The rotated factor pattern details the correlations between factors and variables (Appendix 9). In each column (factor) there are a few high, many low and numerous moderate loadings between variables and factors. This indicates that that several variables will be complex (load on more than one factor) unless a cut-off of .42 is adopted for interpretation. With a cut-off of .42, only two variables load on more than one factor, and 62 (47%) do not load on any factor. Table 5.6.3.6 was generated to assist interpretation, leaving loadings < .42 blank. Full details of all loadings are contained in Appendix 9.

The six factors are interpretable and conceptually relevant to clinical governance climate (table 5.6.3.6). Factor 1 contains items relating to the success of QI initiatives in improving services. These include the balance of external imposition and internal selection of QI activities; the extent to which QI is lost in the day-to-day of service provision; and degree of management support to deliver service changes. This is clearly concerned with the successful delivery of an integrated and planned programme of quality improvement.

Factor 2 contains items related to proactive risk management. It contains items concerned with the extent to which such information informs decision-making, as well as basic items concerned with systematic collection of risk data. Factor 3 contains items related to unjust blame and punishment, also known as the 'blame culture', considered corrosive to quality improvement. Items cover punitive conduct within appraisal systems and interaction between team members as well as the more obvious items relating to investigative procedures during clinical incidents. Factor 4 directly concerns the way that colleagues work together. It includes items on the level of mutual understanding of roles and skills, as well as levels of honesty and respect between colleagues. Factor 5 contains items concerned with training and development opportunities, including the assessment and meeting of development needs, and the availability of training in areas specifically related to clinical governance such as critical appraisal. Finally, factor 6 relates to organisational learning. This factor contains items related to information sharing across organisational boundaries and with others outside the organisation, in order to improve service quality. All sub-scales are positively scored i.e. a high score is indicative of a climate conducive to good clinical governance.

Table 5.6.3.6: Factor loadings and communalities (h²) for Principal Factors extraction and Varimax rotation of CGCQ (factor loadings < .42 omitted)

	Factors						h ²
	1	2	3	4	5	6	
87 lots of QI initiatives; little change	.609						.505
154 few opportunities to use new skills	.552						.550
128 no knowledge of good practice in org.	-.552						.544
138 immediate pressures precedence over QI	-.547						.466
125 no shared service goals	.547						.564
139 QI mainly a response to external pressure	.544						.466
94 First known of QI elsewhere is the effects	.529						.424
56 planning QI lost in day-to-day	-.523						.488
152 react to problems, rather than prevent	-.516						.513
80 clear vision of org.	.510						.496
155 forced into change, rather than encourage	.495						.541
96 service improvement crisis led	.490						.563
122 work in same way as always	-.489						.377
104 no time to share ideas	.486						.437
123 RM largely paper exercise	.481						.469
156 people motivated to make changes	.475						.452
77 no support to deliver changes	.468						.524
99 QI imposed from above	-.468						.385
153 people are motivated to improve quality	.463						.491
110 we don't address accidents waiting	.458						.524
64 good practice isolated in pockets	.442						.544
106 shared vision of service delivery	.434						.425
109 pressure to 'solve' quickly	-.431						.390
141 few opportunities to consolidate learning	-.429						.367
67 risks examined systematically		.661					.603
50 collect info on clinical risks		.636					.466
101 systematically assess risks		.597					.503
83 don't collect info on most important risks		.593					.612
142 risk asst updated following incidents		.570					.474
118 risk info used to inform decisions		.545					.475
97 action follows identified risks		.538					.512
117 risk policies shared org. wide		.523					.374
90 no common RM approach		.499					.364
150 use specific QI techniques		.460					.417
147 failure used as learning opportunity		.452					.526
66 identified risks not addressed		.439					.479
27 Incidents made to feel guilty			.646				.629
34 how happened not who erred			.616				.604
81 atmosphere of blame			.602				.621
1 assumption 'someone to blame'			.601				.524
22 error reporting a stick for clinicians			.544				.476
29 appraisals used to punish staff			.535				.482
53 we look for system error not blame indivs.			.520				.567
32 unsafe to be open with colleagues			.493				.439
42 supported if made mistake			.481				.386

	Factors						h ²
	1	2	3	4	5	6	
127 appraisals a waste of time			.426				.404
54 understand colleague's roles				.603			.536
30 know colleagues skills				.562			.386
146 don't understand roles	.451			.538			.601
124 don't know what colleagues expect of me	.458			.471			.535
151 same standing				-.459			.435
108 respect for contributions of all				.427			.538
59 colleagues are dishonest with each other				.424			.370
38 people own up to mistakes				.424			.321
86 no training in searching for evidence					.521		.420
25 appraisal skills training available					.512		.324
52 help with EBP available					.503		.480
28 development needs match org needs					.458		.424
129 development needs assessed					.441		.452
31 protected QI time					-.429		.397
55 appraisal not identify staff needs					.421		.412
126 time to reflect					-.420		.316
71 share practice issues across org.						.522	.549
10 share ideas outside org.						.476	.410
93 time to disseminate practice						-.473	.637
3 improve own areas, not work together						-.470	.425
91 different teams share good practice						.467	.604
44 work together across teams for QI						.442	.533

Extraction Method: Principal Axis Factoring. Rotation Method: Varimax with Kaiser Normalization.
Cut-off loading = 0.42 (18% shared variance)

Factor 1= A planned and integrated quality improvement programme
Factor 2= Proactive risk management
Factor 3= (absence of) Climate of unjust blame and punishment
Factor 4= Working with colleagues
Factor 5= Training and development opportunities
Factor 6= Organisational learning

5.7 Internal consistency (reliability) and external validity

5.7.1 Reliability and validity

Internal consistency (reliability)

Scale reliability is primarily concerned with consistency. In order to use a scale with confidence, we need to be confident that differences between scale readings when applied to two different subjects reflect genuine differences between the subjects, not inconsistencies in the measurement instrument itself. There are a number of potential approaches to reliability testing, including *test-retest*, *internal consistency*, *split-half* and *parallel-form* (Oppenheim 1992, DeVellis 1991, Wilkin 1992).

(a) Test-retest reliability

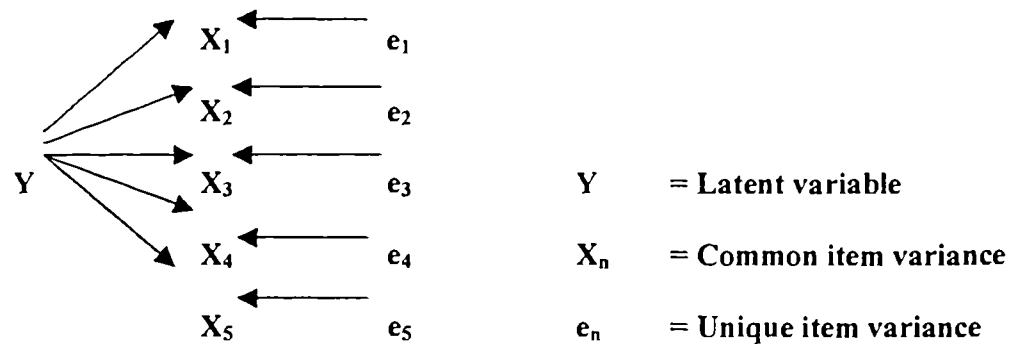
More accurately known as temporal stability, test-retest reliability assesses the ability of a measure to reproduce the same results over multiple administrations to the same sample and under the same conditions within a short period. The reliability of the measure is estimated by the correlation between the sets of results. While this is intuitively appealing, the approach is beset by a number of difficulties. These include resistance to administration, and a 'practice effect' where over repeated administration respondents 'learn' how to fill in the questionnaire. Under such circumstances, the underlying requirement of test-retest reliability to administer the same test under the same conditions is broken. There are also practical requirements of fieldwork, as the procedure requires multiple administrations and collections of data, with the potential for attrition in response rates over

time. Further, temporal stability is only advantageous when it is known that the construct being measured is itself stable over time. Potential confounding factors include change in the construct over time and change in the subjects over time. Finally, recorded stability over time may simply be due to the inability of a measure to detect meaningful change (poor measure sensitivity). Given the problems over such confounding and the difficulties of collection, this study did not utilise test-retest methods.

(b) Internal consistency

Based firmly on classical scaling theory, the internal consistency method defines scale reliability as the proportion of variance attributable to the true score of a *latent variable* (DeVellis 1991). The latent variable is the underlying construct that a scale is intended to reflect. It is not directly observable, and changes over time. Because the latent variable is not directly observable, it is tapped by examining correlations between items in the item pool. Scores on items in the item pool are said to be *caused* by the latent variable (see Box 5.7.1). The notion of a causal relationship between a latent variable and items implies that there is a correlation between the item value and the true score on the latent variable. In other words, each item gives an indication of how strong the latent variable is. As each of the items taps the latent variable, item scores should correlate; the items should tend to vary jointly. By examining the correlation between items purporting to measure a single latent variable, we can infer how strongly each item correlates with the latent variable. This study draws on two principal approaches to internal consistency reliability: principal factor analysis and coefficient alpha.

Box 5.7.1: The relationship between latent variable, indicator items, and unique variance (error)



Each indicator item varies as a function of the source of variance common to itself and other items, caused by the latent variable, and unique, unshared variance, known as error. Unique item variance is a residual, representing all sources of variation not accounted for by other causes, including measurement error. It represents the discrepancy between actual values of X_n and what we would predict it to be if we had knowledge of Y .

Principal factors analysis

As more than one latent variable was anticipated to explain variation between CGCQ items, the number of latent variables (factors) underlying the item pool was examined empirically using factor analysis. The central objective in factor analysis is to understand the nature of the common factors and their relationships to each other and to the items. Despite the fervour of some adherents, factor analysis does not provide mystical insights into the underlying nature of things. A common factor is simply a factor common to more than one item; the items selected for inclusion determine the obtained factor solutions. Thus, the selection and modification of a pool of items has a substantial impact on a factor structure,

either adding or deleting common factors. Since the influence of common factors on items gives rise to correlation between the items, the existence of correlated items indicates the existence of common factors. By looking at correlations between items, it is possible to estimate the number of common factors (latent variables) and obtain numerical coefficients representing the degree of effect of each factor on each item. This facilitates an interpretation of the nature of the common factors. Section 5.6 above details the exploratory factor analysis of the CGCQ item pool.

Coefficient alpha

Internal consistency refers to the consistency of a set of items within a scale (Cramer 1998), and seeks to determine whether the items belong to the same conceptual domain (Bowling 1997). High internal consistency is important as it increases the power of a test i.e. the probability of detecting significant relationships of particular magnitudes (Bruner and Hensel 1993). The internal statistical reliability procedure Cronbach's Alpha (Cronbach 1951) was used to examine items associated with factors. Alpha is defined as the proportion of a scale's total variance attributable to a common source i.e. the true score of a latent variable underlying the items (DeVellis 1991:27). Bland and Altman (1997) suggest Alpha values of over 0.7 are satisfactory. Section 5.7.2 below details CGCQ domain alpha co-efficients.

(c) Split-half consistency

This approach avoids the 'practice effect' identified above, in which respondents can 'learn' how to score well on the instruments. Here, the development sample is split into two sub-

samples, the first used to develop the measure and the second for cross-validation.

Following derivation of an optimal scale on the first sub-sample, the second can be used to attempt to replicate the findings. Replication of findings on the second sample increases confidence in the initial findings. It is always possible that in the first sample unstable, chance factors can be confused with reliable co-variation. As the data in the second sample did not influence item selection, there is no opportunity for attributing reliability to what are in reality chance results.

(d) Parallel-form consistency

This method also avoids the 'practice effect', by a different method to that of split-half consistency. In this approach, two equivalent forms of the scale are developed, identical in every way except for the content of the items. This is particularly useful in experimental or quasi-experimental pre-post designs, as it avoids the necessity for exposing subjects to the same measure twice within a short time period. The problem is that the development of two scales that are equivalent in every way has implications for the scale of fieldwork and testing required, as it is very time intensive. Consequently, the present study will not use parallel-form methods.

(e) Inter-rater reliability

This is an important consideration when interviewers administer measures to subjects and make judgements based on observations. As the scale is self administered, this criterion is inappropriate.

Scale validity

There are a number of ways of expressing validity, but in principle each indicate the degree to which scales measure what they are intended to measure. The study uses three types of validity checking: content; discriminant; and construct validation.

(a) Face validity

This refers to the superficial 'sense' of the indicators, and whether they are easy to understand. On its own, this is a very weak criterion, a basic minimum standard.

(b) Content validity

A considerably more robust construct than face validity, content validity refers to the choice of, and relative importance given to, items included in the item pool. A researcher may make the case for content validity either conceptually or empirically. The former relies on the strength of the conceptual framework of the scale, with reference to relevant literature(s). This draws upon the notion of a community of understanding. The latter empirical approach appeals to involvement of others in the generation and selection of items. The present study makes use of both approaches. Conceptually, the domains of interest are generated from literatures related to clinical governance, discussed in chapter 5. Empirical content validity is demonstrated by the developmental qualitative focus groups and the content validity study of operational clinical governance managers, outlined in sections 5.3 and 5.5 respectively.

(c) Criterion validity

Criterion validity is the ability of an instrument to correspond with a simultaneously collected pre-existent measure (concurrent criterion validity) or to forecast a future value (predictive criterion validity) (Bowling 1997). This is particularly important in the design of new measures, in order to corroborate that they accurately measure the phenomenon of interest. However, this assumes that well validated alternative measures are available. As the lack of a 'gold standard' measure of clinical governance climate was one of the main reasons for developing the CGCQ, it was not possible to test its criterion validity against such a measure. In such circumstances recourse is usually made to internal consistency methods of reliability (Oppenheim 1991). The assumption is that if a scale is internally consistent, it is at least consistently measuring an underlying construct. The present study supplemented internal consistency with an assessment of the discriminant and construct validity of the factor solution.

(d) Discriminant validity

Discriminant validity is the ability of a measure to detect cross-sectional differences between subjects when such differences on the construct are either expected or known. Given the variety of organisational context, one would expect to find significant differences in sub-scale scores between organisations. Indeed, if the instrument were not sensitive enough to detect such differences, it would not be worthwhile. Similarly, within any organisation one would expect to find differences in sub-scale scores between professional groups. For example, across all organisations one would expect clinical groups to report higher levels of 'unjust blame' than managerial groups, given their relative positions and

responsibilities. Section 5.7.3 below summarises the ability of CGCQ sub-scales to detect such expected differences within the sample.

(e) Construct validity

Where there is no reference gold standard against which to explore criterion validity, empirical evidence may support the inference that a measure has meaning (Wilkin 1992). Such an approach typically includes comparisons with related measures, and predictions about the distributions of scores among subgroups of respondents. Construct validity is the ability of an instrument to confirm expected hypotheses, and is assessed by exploring expected relationships between factor scores and other constructs given an initial hypothesis.

There is potential for correlation between some of the domains identified in the new clinical governance measure and other measures. A balance was struck between inclusion of enough extra instruments to conduct a robust construct validation, and the addition of as few extra instruments as possible, in order to increase potential response rates. Two short and well-validated measures were chosen: the Competing Values Questionnaire (CVQ) and the Maslach Burnout Inventory (MBI). Details of these measures and of the hypotheses forming the basis of the construct validation of the CGCQ factor solution, together with results, are reported in section 5.7.3 below.

5.7.2 CGCQ Internal reliability (Cronbach's Alpha)

Scores of internal scale consistency show high reliability for all six sub-scales (table 5.7.2.1), with Cronbach's α ranging from 0.81 to 0.95 (Nunnally 1978). To minimise item redundancy, items were removed if they contributed less than 0.01 to scale reliability. This resulted in the exclusion of three indicators from factor 1, two indicators from factor 4 and one indicator each from factors 2, 3, 5 & 6. These reliability measures suggest a high degree of internal consistency for the six factors.

Table 5.7.2.1: Cronbach's alpha coefficients for CGCQ sub-scales

Items and α values for scales	
Factor 1: A planned and integrated QI programme ($\alpha = .95$)	21 items
87	There are lots of quality improvement initiatives, but little change
154	There are few opportunities to use new skills learned as part of development
128	People don't know about good practice taking place in other parts of the organisation
138	Immediate pressures are always more important than quality improvement
125	People don't seem to have shared service goals
139	Quality improvement activity is largely a response to external pressure
94	The first we know of quality improvements elsewhere in the organisation is when we feel the effects
56	Long-term planning for quality improvement gets lost in the day-to-day
152	We react to problems, rather than try to prevent them
80	There is no clear vision of what the organisation is trying to achieve
155	People are forced into making service changes, rather than encouraged to make them
96	Around here, service improvements tend to be crisis led
104	There is no time to get together to share ideas
156	People are highly motivated to make changes to clinical practice
77	There is no support to deliver service changes
99	Quality improvement is imposed from above rather than built from below
153	People are motivated to improve quality
110	We don't address the accidents waiting to happen
64	Good practice stays isolated in pockets
106	People share a common vision of service delivery
109	There is pressure to 'solve' problems quickly rather than take time and do it properly

Items and α values for scales

Factor 2: Proactive risk management ($\alpha = .90$) 11 items

- 67 Clinical risks are examined systematically
- 50 We collect information on clinical risks
- 101 We systematically assess clinical risks
- 83 We don't collect information on the clinical risks that matter most
- 142 Risk assessment processes are updated in the light of clinical incidents
- 118 Clinical risk information is used routinely to inform decisions
- 97 When a clinical risk is identified, there is action to address it
- 117 Clinical risk policies are shared throughout the organisation
- 90 There is no common approach to risk management
- 147 When something fails, it is used as a learning opportunity
- 66 Identified risks simply remain unaddressed

Factor 3: Climate of blame and punishment ($\alpha = .90$) 9 items

- 27 People involved in clinical incidents are made to feel guilty
- 34 The emphasis is on how an incident happened not who made the mistake
- 81 We work in an atmosphere of blame
- 1 When things go wrong there is an automatic assumption that 'someone is to blame'
- 22 Error reporting systems are basically a stick to beat clinicians with
- 29 Staff appraisals are used to punish staff
- 53 When there is an error, we look for failure in systems, rather than blame individuals
- 32 It is unsafe to be open and honest with colleagues
- 42 People who make mistakes are supported

Factor 4: Working with colleagues ($\alpha = .85$) 6 items

- 30 People have a good knowledge of the skills of their colleagues
- 146 Colleagues don't seem to understand each other's roles
- 124 People don't know what their colleagues expect of them
- 151 Everyone has the same standing, regardless of professional background
- 108 There is mutual respect for everyone's contribution
- 59 Colleagues are dishonest with each other

Factor 5: Training and development opportunities ($\alpha = .81$) 8 items

- 86 There is no training available in searching for research evidence
- 25 Critical appraisal skills training is available to those who want it
- 52 Technical help with Evidence Based Practice is available
- 28 Career development needs are addressed alongside the strategic needs of the service
- 129 Development needs are regularly assessed
- 31 We have protected time for quality improvement activity
- 55 Appraisal does not identify the real development needs of staff
- 126 There is time to reflect on practice

Factor 6: Organisational learning ($\alpha = .87$) 5 items

- 71 People share practice issues with others in different parts of the organisation
 - 10 Good practice ideas are shared with others outside the organisation
 - 93 People devote time to disseminating good practice
 - 91 Teams from different parts of the organisation share their good practice
 - 44 We work together across teams to make quality improvements
-

5.7.3 External validity

While exploratory factor analysis is useful for reducing indicator item pools and describing a latent factor structure, it is not sufficient to establish the external validity of identified factors. An additional preliminary assessment of the external validity of the six-factor CGCQ model was undertaken by assessing its ability to discriminate between subgroups (*discriminant validity*) and conformity to expectations (*construct validity*). The results provide support for the validity of at least five of the six factors derived from the factor analysis.

5.7.3.1 Discriminant validity

Results of the evaluation of discriminant validity are summarised in table 5.7.3.1.1. Factors reflected anticipated differences as a function of trust (figure 5.7.3.1.1) and professional background (figure 5.7.3.1.2). In the former analysis, results were based on data from the four trust sites only, excluding the additional 171 cases taken from HSMC conference attendees used to boost the sample size. Using simple summated scores, the CGCQ discriminated between trusts at $p < .01$ for all factors other than factor 4, working relationships. The analysis of professional group included all cases, and showed that using simple summated scores, the CGCQ discriminated between professional groups at $p < .01$ on all factors other than factor 4, working relationships, and factor 6, organisational learning. These differences are explored further using uncorrelated regression factor scores to avoid exaggerating effects in a multiple analysis of variance (MANOVA) in section 6.2 below.

Table 5.7.3.1.1: Discriminant validity: differences and similarities in CGCQ factor scores by profession and trust

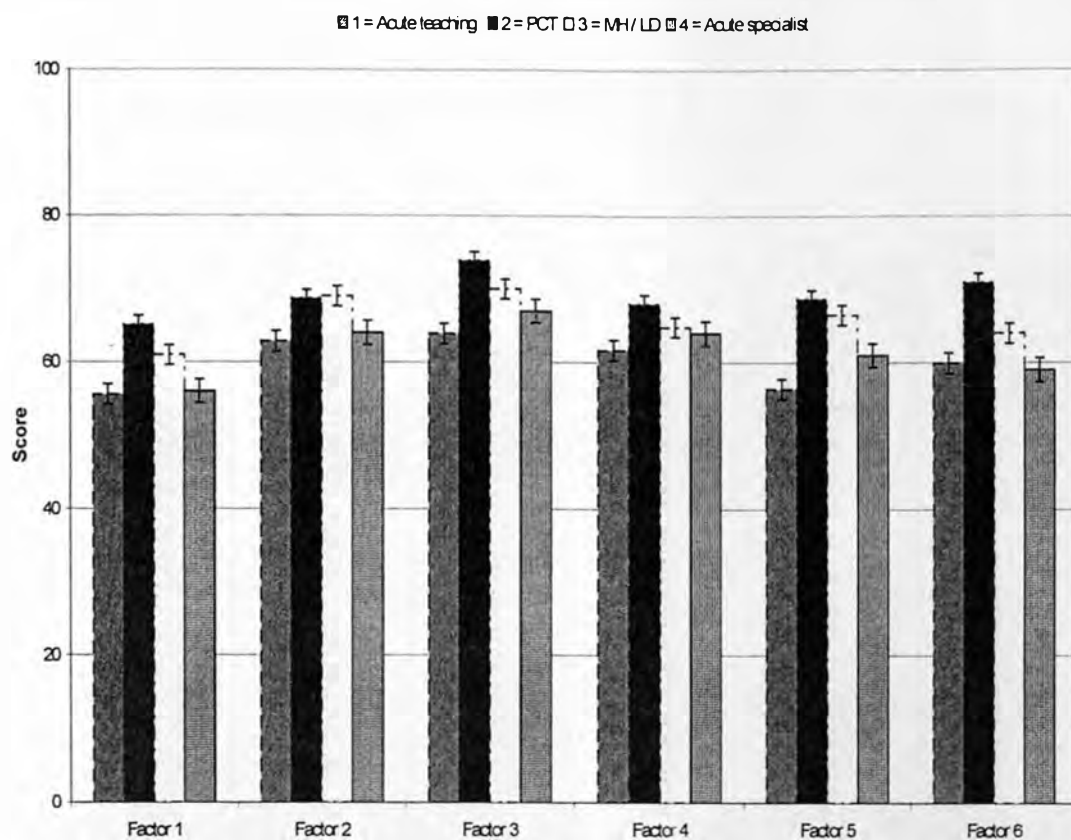
Independent variables	CGCQ Factors					
	Planned and integrated QI	Pro-active risk management	Absence of unjust blame	Working with Colleagues	Training & development	Organisational learning
Profession	.008**	.000***	.000***	.815	.002**	.152
Trust	.000***	.003**	.002**	.071	.000***	.000***

*** = Differences in factor scores between subgroups of independent variable $p < .001$

** =Differences in factor scores between subgroups of independent variable $p < .01$

Raw CGCQ item scores summed and standardised to produce six factor scores of 0-100. One-way analysis of variance (F -test) used to compare mean differences in CGCQ factor scores across sub-groups of independent variables. Where differences in factor scores across sub-groups are noted, the lower the p -value the better the discrimination.

Figure 5.7.3.1.1: Inter-trust comparison: mean factor scores

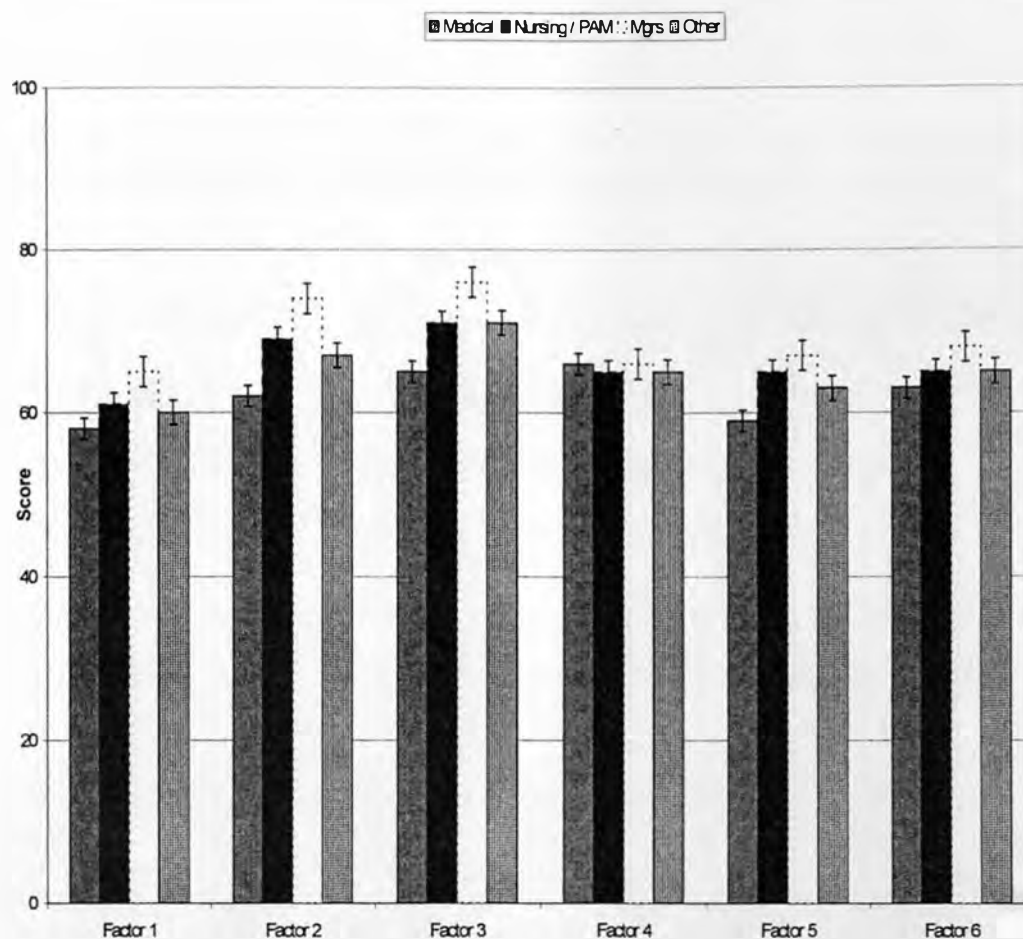


		Quality improvement	Risk management	Absence of blame	Positive relationships	Training and development	Organisational learning
Sample	Mean (SD)	60 (13)	67 (12)	69 (14)	65 (14)	64 (13)	64 (14)
n = 235	95% CI	58 - 62	65 - 68	68 - 71	63 - 67	62 - 65	62 - 66
1	Mean (SD)	56 (13)	63 (15)	64 (17)	62 (13)	56 (15)	60 (16)
(n= 50)	95% CI	52 - 60	59 - 62	59 - 69	58 - 65	52 - 61	55 - 65
2	Mean (SD)	65 (12)	69 (10)	74 (13)	68 (13)	69 (11)	71 (13)
(n= 71)	95% CI	62 - 68	66 - 71	71 - 77	65 - 71	66 - 71	68 - 74
3	Mean (SD)	61 (12)	70 (11)	71 (14)	65 (14)	66 (11)	64 (13)
(n= 68)	95% CI	58 - 64	67 - 72	67 - 74	61 - 68	64 - 69	61 - 67
4	Mean (SD)	56 (11)	64 (10)	67 (12)	64 (13)	61 (11)	59 (12)
(n= 46)	95% CI	53 - 59	61 - 66	65 - 72	60 - 67	58 - 65	55 - 62
anova	F	8.108	4.887	593	2.379	11.402	10.033
	Sig.	.000***	.003**	.002**	.071	.000***	.000***

*** = Differences in factor scores between trusts significant at $p < .001$ level

** = Differences in factor scores between trusts significant at $p < .01$ level

Figure 5.7.3.1.2: Profession-specific assessment of climate: mean factor scores



		Quality improvement	Risk management	Absence of blame	Positive relationships	Training and development	Organisational learning
<i>Sample</i>	<i>Mean (SD)</i>	62 (13)	69 (13)	72 (14)	65 (13)	65 (13)	66 (15)
(n = 403)	<i>95% CI</i>	60 - 63	68 - 70	71 - 74	64 - 67	63 - 66	64 - 67
Medical	<i>Mean (SD)</i>	58 (15)	62 (16)	65 (17)	66 (13)	59 (15)	63 (16)
(n = 58)	<i>95% CI</i>	54 - 62	58 - 66	61 - 70	63 - 69	55 - 63	59 - 67
Nursing	<i>Mean (SD)</i>	61 (12)	69 (10)	71 (14)	65 (14)	65 (12)	65 (15)
(n = 151)	<i>95% CI</i>	59 - 63	67 - 70	69 - 73	63 - 67	63 - 67	62 - 67
Mgrs	<i>Mean (SD)</i>	65 (13)	74 (12)	78 (13)	66 (13)	67 (12)	68 (15)
(n = 129)	<i>95% CI</i>	62 - 67	72 - 76	75 - 80	64 - 68	64 - 69	65 - 70
Other	<i>Mean (SD)</i>	60 (14)	67 (12)	71 (12)	65 (14)	63 (11)	65 (16)
(n = 65)	<i>95% CI</i>	56 - 63	64 - 70	68 - 74	61 - 68	60 - 68	61 - 69
anova	F	3.981	14.079	11.995	.314	4.947	1.77
	Sig.	.008**	.000***	.000***	.815	.002**	.152

Note: missing data (professional group not specified) in 3 cases

*** = Differences in factor scores between professional groups significant at P< .001 level

** = Differences in factor scores between professional groups significant at P< .01 level

5.7.3.2 Construct validity

(i) Competing Values Questionnaire (CVQ)

The competing values model of organisational culture (Quinn 1988; Quinn and Rohrbaugh 1981, 1983) seeks to explain differences in the values underlying organisational effectiveness models. The model postulates that while cultural artefacts such as traditions and myths are organisation-specific, their underlying values are generic. As organisational cultures are the result of varying emphases on a limited number of values prevalent in society, they are therefore amenable to quantitative summary. Quinn and Kimberly (1984) employed the model to explore deep structures of organisational culture along two sets of continua: focus (individual Vs organisational); and structure (stability Vs change).

The continua outlined above combine to describe four 'ideal' cultural types. Based on norms of affiliation, group (clan) cultures seek to secure compliance with organisational mandates through trust, tradition and member commitment. They emphasise the development of potential, value participation in decision-making, and implement change through consensus building. Predicated upon change, developmental (open) cultures seek to motivate individuals through appeals to the importance of required tasks, emphasising growth and acquisition. In contrast, hierarchical cultures reflect the values and norms of bureaucracy. They assume stability, seeking to ensure compliance with mandates via formal rules enforced by sanctions and regulations. Finally, rational cultures assume achievement through planning, productivity and efficiency, seeking to motivate individuals through the belief that competency leads to rewards. Real organisations contain

combinations of such values leading to idiosyncrasies in the myths, symbols and language of organisational life. The CVQ operationalises the competing values model, requiring respondents to indicate the extent to which their organisation reflects characteristics associated with each of the four ideal types along four institutional dimensions: character, leader, cohesion and emphasis.

(ii) The Maslach Burnout Inventory -General Survey (MBI-GS)

The MBI-GS measures respondents' relationships with their work on a continuum from engagement to burnout, and has been widely applied in health care settings (Koivula *et. al.* 2000; Stordeur *et. al.* 2001; Taormina and Law 2000). Engagement is an energetic state in which workers pursue excellent performance and have confidence in their effectiveness. Conversely, burnout is a state of exhaustion in which a worker is cynical of the value of their work and doubtful of their effectiveness. The MBI-GS consists of three sub-scales: exhaustion, cynicism and professional efficacy (Schaufeli, Leiter and Kalimo 1995). Exhaustion items consist of references to fatigue and emotional draining, cynicism items reflect indifference or a distant attitude towards work, and professional efficacy focuses on accomplishments and expectations. High scores on exhaustion and cynicism and low scores on professional efficacy reflect a high degree of burnout. Cynicism reflects dysfunctional coping, as employees develop cynicism about work in order to gain distance from exhausting demands. As cynicism reduces energy and creative solutions to work problems, it correlates positively with exhaustion and negatively with professional efficacy.

Assessment of the construct validity of CVQ and MBI-GS within the current dataset

In the present sample, examining within-scale correlations for each measure (table 5.7.3.2.1) assessed CVQ and MBI-GS construct validity. For the CVQ, a competing values profile was devised for each individual by summing ratings for each cultural type across the four dimensions of institutional characteristics, leadership, cohesion and emphases and dividing by four. This produced a score for group, developmental, hierarchical and rational culture for each individual. Overall institutional profiles for the four cultural types for each organisation were obtained by averaging each of the four scores across respondents from the organisation. Scoring on the MBI-GS involved computing an average rating for each individual across the items within each of the three sub-scales exhaustion, professional efficacy and cynicism.

In accordance with the competing values model, scores for group (clan) culture had high negative correlations with scores for rational culture, and scores for hierarchical structure had high negative correlations with scores for developmental culture (table 5.7.3.2.1). On the MBI-GS, scores for cynicism correlated positively with exhaustion and negatively with professional efficacy as expected (table 5.7.3.2.1). Both the pattern and magnitude of the relationships supports the construct validity of the CVQ and MBI-GS in the present sample, and the validity of aggregating data to the organisational level in subsequent analysis.

Table 5.7.3.2.1: Intra-scale correlation: CVQ culture-type and MBI sub-scales (Pearson's r)

CVQ	Group	Developmental	Hierarchical	Rational
Group	1			
Developmental	.286***	1		
Hierarchical	-.492***	-.595***	1	
Rational	-.524***	-.454***	.039 (NS)	1

MBI-GS	Exhaustion	Professional efficacy	Cynicism
Exhaustion	1		
Professional efficacy	-.271***	1	
Cynicism	.493***	-.519***	1

*** = $p < .001$

As hypothesised, the CVQ shows strong negative associations between group (clan) & rational, and hierarchical & developmental scores. In addition, as hypothesised MBI-GS cynicism is strongly positive associated with exhaustion and strongly negatively associated with professional efficacy. The pattern and magnitude of associations supports the construct validity of CVQ and MBI-GS scales in the present sample, and validates the use of aggregated data in organisational level analysis

Hypothesised relationships between CGCQ sub-scales and CVQ / MBI-GS

CGCQ construct validation hypothesised the direction of correlations between CGCQ sub-scales and the Competing Values Questionnaire (CVQ) and the Maslach Burnout Inventory – General Survey (MBI-GS). The hypotheses assume positive scoring of CGCQ sub-scales i.e. the higher the sub-scales score the more positive the climate for clinical governance.

Hypothesis 1: CGCQ sub-scale scores positively correlate with scores on 'group' and 'developmental' domains of the CVQ. The building of trust, member commitment and consensus around change and growth measured by the CVQ is expected to be associated with a climate conducive to engagement with the long-term clinical governance agenda of quality improvement (Shortell *et. al.* 1995; Halligan and Donaldson 2001).

Hypothesis 2: CGCQ sub-scale scores negatively correlate with scores on the 'hierarchical' domain of the CVQ. In contrast to hypothesis 1 above, the existence of a strong hierarchical element to organisational culture is associated with a climate less conducive to active engagement with quality improvement (Shortell *et. al.* 1995), and therefore the long-term clinical governance agenda (Davies and Mannion 1999). Hierarchical (bureaucratic) modes of governance seek to enforce mandates formally via sanctions and regulations, rather than seeking the active consent of professional groups (Hood 1996).

Hypothesis 3: CGCQ sub-scale scores positively correlate with scores on the 'professional efficacy' domain of the MBI-GS. The assumption is that professionals engaged in the long-term clinical governance agenda are likely to feel positive about their professional abilities; and those that feel positive about their abilities are more likely to be motivated to make quality improvements.

Hypothesis 4: CGCQ sub-scale scores negatively correlate with scores on 'exhaustion' and 'cynicism' domains of the MBI-GS. In contrast to hypothesis 3 above, excessive demands on staff time is likely to be associated with a climate in which the long-term agenda of clinical governance is unable to flourish. One would also expect cynicism to arise under such circumstances.

All four hypotheses concerning the behaviour of CGCQ sub-scales in relation to CVQ and MBI-GS sub-scales were confirmed (table 5.7.3.2.2). CVQ group (clan) & developmental cultures are positively associated, and hierarchical cultures negatively associated, with CGCQ sub-scales. This is particularly so in relation to CGCQ sub-scales for planned & integrated QI, training & development, and organisational learning. Additionally, MBI-GS professional efficacy is positively associated, and exhaustion & cynicism negatively associated, with CGCQ sub-scales. This suggests that organisations that promote growth, development, trust and commitment are likely to foster quality improvement activity, proactive risk management and organisation learning. Conversely, the negative correlation between 'hierarchical' CVQ scores and CGCQ factors suggests that the pursuit of compliance with mandates via formal rules enforced by sanctions promotes a 'blame' focus

while hampering quality improvement and organisational learning. This is as expected, adding weight to the external validity of the CGCQ factors.

Table 5.7.3.2.2: Concurrent validity: correlation of CGCQ with CVQ and MBI sub-scales (Pearson's r)

	CGCQ sub-scales					
	Planned and integrated QI	Pro-active risk management	Absence of unjust blame	Working with Colleagues	Training & development	Organisational learning
CVQ sub-scales:						
Group	.309***	.253***	.251***	.316***	.302***	.276***
Developmental	.539***	.381***	.434***	.414***	.484***	.490***
Hierarchical	-.459***	-.381***	-.376***	-.381***	-.369***	-.401***
Rational	-.286***	-.180***	-.214***	-.295***	-.272***	-.280***
MBI sub-scales:						
Exhaustion	-.446***	-.274***	-.394***	-.337***	-.355***	-.277***
Professional efficacy	.350***	.273***	.234***	.339***	.279***	.273***
Cynicism	-.507***	-.372***	-.392***	-.443***	-.350***	-.327***

*** = $p < .001$ Raw CGCQ item scores summed and standardised to produce six factor scores of 0-100. Higher correlations reflect stronger associations. As hypothesised, CVQ group and developmental cultures are positively associated, and hierarchical cultures negatively associated, with CGCQ sub-scales. This is especially so for quality improvement and organisational learning. In addition, as hypothesised MBI-GS professional efficacy is positively associated, and exhaustion and cynicism negatively associated, with CGCQ sub-scale scores.

5.8 Discussion

The purpose of the study was to derive a measure of clinical governance climate in a population of health care trust staff. Results show that it is possible to develop a valid survey instrument of clinical governance climate. Data met criteria of internal consistency, predictable relationships with other phenomena and discrimination between professional staff groups and organisations. While exploratory factor analysis is useful for reducing indicator item pools and describing a latent factor structure, it is not sufficient to establish the external validity of identified factors. An additional preliminary assessment of the external validity of the six-factor CGCQ model was undertaken by assessing its ability to discriminate between subgroups (*discriminant validity*) and conformity to expectations (*construct validity*). The results provide support for the validity of the factors derived from the factor analysis.

In terms of discriminant validity, simple summated CGCQ factor scores identified significant ($p < .001$) differences between both trusts and professional groups. In an assessment of concurrent construct validity, statistically significant and moderate correlations were observed between factor scores and sub-scales of the Maslach Burnout Inventory (MBI) and Competing Values Questionnaire (CVQ). The moderate negative correlations between exhaustion and cynicism factors on the MBI and the CGCQ factors are as expected. The more exhausted and cynical the workforce, the less likely progress in quality improvement, proactive risk management and organisational learning becomes. Positive correlations between professional efficacy on the MBI and CGCQ factors reflects

the intuitive assumption that a workforce that believes in its ability to make a difference is likely to make more progress than one that doubts its efficacy.

Similarly, correlations between CVQ and CGCQ factors were as expected. Positive correlations between 'developmental' and 'group' CVQ factors and CGCQ factors suggests that organisations which promote growth, development, trust and commitment are likely to foster quality improvement activity, pro-active risk management and organisation learning. Conversely, the negative correlation between 'hierarchical' CVQ scores and CGCQ factors suggests that the pursuit of compliance with mandates via formal rules enforced by sanctions promotes a 'blame' focus while hampering quality improvement and organisational learning. This is as expected, adding weight to the external validity of the CGCQ factors.

There are a number of limitations to the power and generalisability of the findings. Firstly, an exploratory factor analysis of a 157-item questionnaire on a sample of 406 respondents requires cautious evaluation. Problems of response bias may further limit the findings, as it is possible that staff with dramatically more negative (or even positive) experience of clinical governance climate elected not to participate. However, while the lack of accepted 'gold standard' alternative measures made it difficult to assess criterion validity, assessments of discriminant and construct validity both provided support for the external validity of the measure in the present study population. The results show the six-factor model of the CGCQ to be a useful step towards the design and development of such a 'gold standard' instrument.

The negative association between all domains of clinical governance climate on the CGCQ and hierarchical culture on the CVQ (table 5.7.3.2.3) sounds a note of caution for all concerned with the long-term clinical governance agenda within organisations. A reliance on mandates enforced by regulations and sanctions was associated with less quality improvement activity; less learning from others; less pro-active risk management; more unjust blaming; poorer working relationships; and less staff development. The danger is that an overly bureaucratic performance measurement approach that over-emphasises the extent to which organisations are subject to rational decision-making may paradoxically undermine the conditions required for quality improvement; it may be a 'fatal remedy' (Sieber 1981).

Trusts face a difficult balancing act as they seek to align central and local agendas for quality improvement dynamically, while meeting their performance measurement requirements (Degeling, Hunter and Dowswell 2001; Lawton, McKevitt and Millar 2000). One way of reducing the unintended negative consequences of performance management systems is to become reflexively aware of their potential impact and seek methods of redress (Power 2001). This is particularly important in the light of recent developments at the Commission for Health Improvement and its new role as a Commission for Health care Audit and Inspection (CHAI). Use of measures such as the CGCQ will provide trusts with valuable information on staff perceptions on clinical governance climate, which they may use to identify priorities for organisational development activity and assess their efficacy over time. Currently, the danger is that given the policy context, the pursuit of the long-

term agenda may leave organisations open to accusations of not having robust enough performance management. Use of the CGCQ may provide trusts with valuable evidence of a robust, long-term focus on quality improvement, with which to reassure external agencies.

Conclusions

The CGCQ was constructed using focus groups to explore domains of interest identified in clinical governance literature from the perspective of clinicians and managers, in order to develop a draft item pool. A postal validation study assessed the face validity of the items, and an exploratory principal factor analysis explored the latent factor structure of the item pool. Results show that it is possible to develop a valid survey instrument of clinical governance climate. Data met criteria of internal consistency, predictable relationships with other phenomena and discrimination between professional staff groups and / or organisations. The CGCQ thus has known validity and reliability and may be used to assess clinical governance climate.

CHAPTER 6

MEASURING CLINICAL GOVERNANCE CLIMATE: INTER-ORGANISATIONAL AND INTER-PROFESSIONAL EFFECTS

Introduction

The previous chapter explored the development, validity and reliability testing of the Clinical Governance Climate Questionnaire (CGCQ). This chapter reports an application of the CGCQ, drawing attention to the methodological problem of multiple professional sub-cultures through an analysis of the effects of organisational context and professional group on experience of clinical governance climate in four health care trusts.

The cultural complexity of health care organisations was considered in chapter four. Given the variety of professional groups, divisions and teams that work within them, health care organisations are essentially multi-cultural (Davies, Nutley and Mannion 2000; Degeling, Kennedy and Hill 2001). Traditionally, clinicians are socialised into a professional culture in which treatment decisions require professional discretion in the treatment of individual patients. In contrast, the professional culture of managers tends to focus on longer-term planning for whole populations (Ferlie and Shortell 2001). Managers face the difficult task of exploring such differences and securing consensus around the need for multi-disciplinary approaches.

In the context of clinical governance climate, one might reasonably expect to find differences in domain scores both *between* and *within* organisations: the former due to

broader organisational contexts, and the latter to the different experience of professional subgroups. While the previous chapter showed that simple summated CGCQ scores were sensitive enough to detect differences between different organisations and professional groups, this section explores their relative importance and interaction by means of a multivariate analysis of variance (MANOVA).

Analysis explored the impact of professional group, organisation and their interactions on CGCQ domain scores in order to address the following questions:

1. Are differences in clinical governance climate scores across all domains associated with professional group; or organisation; or both? Is there any interaction effect? i.e. is the pattern of responses in different trusts the same for all professional groups, and vice versa?
2. If there is a significant multivariate effect which of the six CGCQ domains does professional group or trust affect? Which professional groups or trusts are significantly different to each other?

6.1 Methodology

The study employed a structured postal questionnaire to explore the clinical governance climate, and was based on data collected for the factor analysis of the CGCQ reported in the previous chapter. Briefly, the sample was drawn from four health care trusts: an acute teaching hospital; a Primary Care Trust (PCT); a community, mental health and learning

disabilities trust; and a specialist acute hospital. An additional 171 respondents were recruited from recent attendees to conferences at the Health Services Management Centre, in order to boost the sample size to 406 in total.

Sampling

Each of the four trust sites drew a random sample of 210 staff from HR records, distributing questionnaires through internal mail systems. In order to boost sample size, an additional sample of 300 healthcare managers was drawn from recent conference attendees at the Health Services Management Centre, University of Birmingham, who received questionnaires mailed to their place of work. In total, respondents returned 406 completed questionnaires directly to the researchers; data was available from 406 subjects with no missing values for indicator items. Given the existence of strong professional subcultures in health care settings (Degeling *et. al.* 1998; 2003), it was important that respondents' professional group was identified in order to facilitate sub-group analysis alongside whole organisation analysis.

Study measures

Detailed in the previous chapter, the Clinical Governance Climate Questionnaire (CGCQ) was used to assess respondent's experience of work in six domains related to clinical governance. Regression factor scores computed under SPSS FACTOR were used in the analysis, due to their normal distribution and lack of correlation. The former is required for small to moderate sample sizes in multiple analysis of variance (MANOVA), and the latter makes interpretation more robust, as significant effects on dependent variables are not

inflated due to the variance that they share with other dependent variables (Tabachnick and Fidell 2001). Scores were standardised between 0-100 for ease of comparison using the mathematical transformation:

$$\text{New score} = (\text{old score} + \text{lowest score}) / (\text{highest score} + \text{lowest score}) * 100$$

The study additionally utilised measures of trust and professional group. The first consisted simply of the respondent's organisation, and professional group was recorded as either managerial, nursing / Profession Allied to Medicine (PAM), medical, or other. A measure of organisational culture, the Competing Values Questionnaire, was also taken for construct validation purposes as reported in the previous chapter, and for descriptive purposes to explore differences between the study sites.

Statistical methods

To explore the effects of professional groups and organisation on CGCQ domain scores, a 4 x 4 factorial design multiple analysis of variance (MANOVA) analysis was performed, using factor scores on the six CGCQ domains as dependent variables. In order to assess the assumptions of normal distribution and homogeneity of variances required by MANOVA, one-sample Kolmogorov-Smirnov tests were used to assess departure from normal distribution of dependent variables, and Box's-M was used to assess the degree of homogeneity of variance. The full factorial model was assessed, including interactions between the independent variables. Wilkes' lambda and Pillai's trace were used to assess the main effect of independent variables on combined independent variables. Significant

main effects were explored further using univariate F-tests with Bonferroni correction for multiple tests. MANOVA was chosen in order to explore interaction effects between the two independent variables i.e. to assess whether the effects of professional group were the same in all trusts; and whether the effects of a trust were the same for all professional groups. One way analysis of variance (ANOVA) was used to assess the significance of differences in mean scores between study sites on each of the four CVQ culture-types.

Testing assumptions: normal distribution and homogeneity of variance

Normal distribution of dependent variables

Dependent variables were regression factor scores of the CGCQ domains taken from SPSS FACTOR in the factor analysis reported above, which automatically produces normally distributed, uncorrelated factor scores.

Homogeneity of variance-covariance

MANOVA assumes that variances between cells in the analysis are from the sample population and can be pooled to create a summary statistic, tested by Box's-M. Box's-M was statistically significant ($F = 1.292$; d.f. 21310.050; $p = .001$), which suggests that the assumption was violated and that results should be considered tentative.

6.2 Results

6.2.1 Descriptive detail on the study sites

Data was gathered on respondents from four trusts: an acute teaching hospital; a primary care trust (PCT); a community/mental health/learning disabilities (C/MH/LD) trust; and an acute specialist hospital.

(i) Acute teaching hospital

Trust 1 is an acute teaching hospital with 934 beds and a total income of £228.4M, operating across 3 sites. It employs over 5270 FTE staff, including 547 medical and over 2660 nursing staff. Services are organised within four directorates, and a wide range of services are provided, including A&E; plastic surgery; cardiac, renal and respiratory medicine; general and transplant surgery; psychiatry; orthopaedics; diabetes; oncology; and palliative care.

(ii) Primary Care Trust (PCT)

Trust 2 became a PCT in April 2000, providing community, mental health and primary care services from 8 locations. It has 316 beds, and is organised around seven service directorates: community dentistry; disabilities and elderly services; occupational therapy; speech and language therapy; community mental health; drug and alcohol abuse; and primary care. It employs over 830 staff, including 44 medical and dental and over 400 nursing staff.

(iii) Community/mental health/learning disabilities trust

Trust 3 is a community/mental health/learning disabilities trust, organised around three directorates. Services include adult and child mental health; drug and alcohol;

physiotherapy; and rehabilitation. It employs over 1160 staff, including 45 medical and dental and over 580 nursing staff. This trust subsequently merged with two local Primary Care Groups (PCGs) in April 2002.

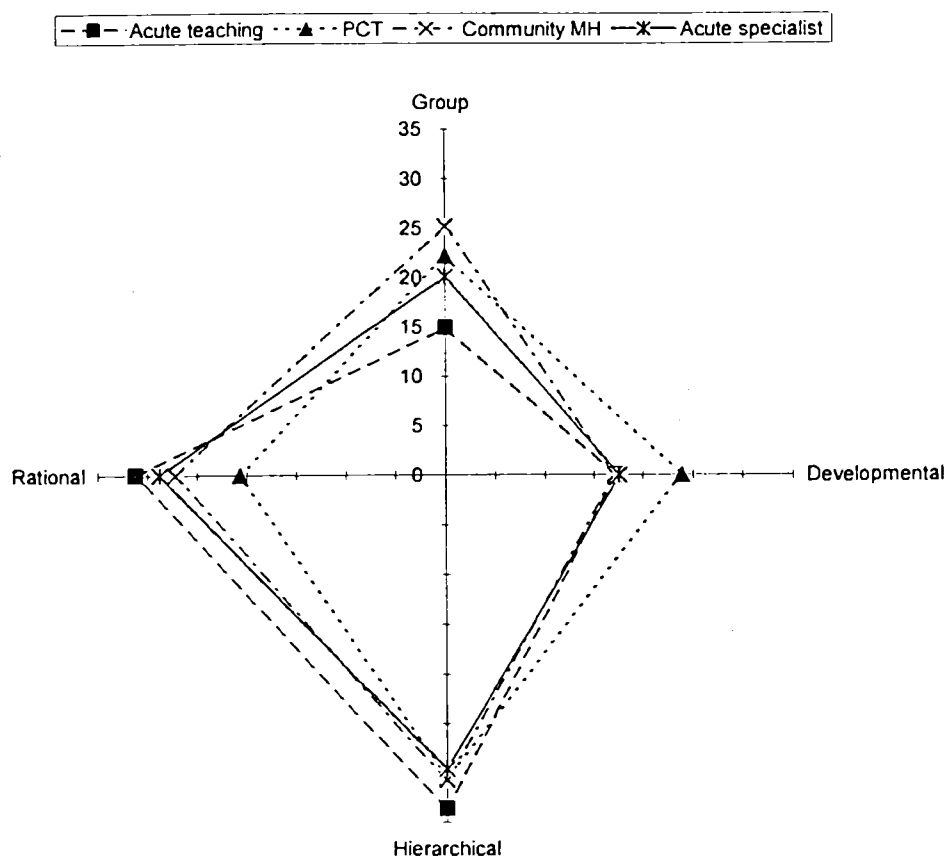
(iv) Acute specialist trust

Trust 4 is a children's hospital with 232 beds and an annual income in excess of £90M, operating across 4 sites. Services are organised into 5 service directorates: child and adolescent mental health; diagnostic and medical support; neuroscience; surgical; and special medical services. Services provided include mental health; neurology; transplant surgery; dermatology; diabetes; oncology; renal; and respiratory medicine.

Competing Values Questionnaire scores by site

A one way analysis of variance was used to assess differences in mean scores between study sites on each of the CVQ culture-types. Significant differences were found between study sites for CVQ clan [group] ($F = 4.244$, 3 d.f., $p = .006$) and CVQ rational ($F = 6.051$, 3 d.f., $p = .001$) culture scores. Post hoc Scheffe tests were used to explore these significant differences further. The overall significant difference for clan [group] culture was due to a significant difference ($p = .008$) between scores for the C/MH/LD trust (mean = 25.24) and the acute teaching hospital (14.88). The overall significant difference for CVQ rational was due to a significant difference between scores for the PCT (Mean = 20.74) and both the acute teaching hospital ($p = .001$, mean = 31.16) and the acute specialist hospital ($p = .034$, mean = 28.78).

Figure 6.2.1.1: CVQ culture sub-scale scores by trust



The PCT had relatively high developmental & group attributes and relatively low rational culture attributes. The acute teaching hospital had relatively high rational & hierarchical attributes, and relatively low group attributes.

6.2.2 Analytical results

1. Are differences in clinical governance climate scores across domains associated with organisation or professional group or both? Is there any interaction effect? i.e. is the pattern of responses in different trusts the same for all professional groups, and vice versa?

MANOVA of trust and profession on CGCQ regression factor scores

Main effects

Multivariate tests identified statistically significant main effects for both trust ($F = 1.953$, $d.f. = 24$, $p = .004$) and profession ($F = 2.126$, $d.f. = 18$, $p = .004$). There was no significant interaction effect between the two independent variables ($F = .932$, $d.f. = 72$, $p = .641$), which means that professional differences operated in the same way across trusts, and trust differences operated in the same way across professional groups.

Table 6.2.2.1: Multivariate tests of significance (Pillai's trace and Wilkes' lambda)

Effect	Test	Value	d.f.	Error d.f	F	Power	Sig.
TRUST	Pillai's trace	.119	24.000	1524.000	1.953	.995	.004
	Wilkes' lambda	.884	24.000	1319.894	1.970	.986	.004
PROF	Pillai's trace	.097	18.000	1140.000	2.126	.988	.004
	Wilkes' lambda	.904	18.000	1069.000	2.160	.984	.003
TRUST*	Pillai's trace	.170	72.000	2298.000	.932	.993	.641
PROF	Wilkes' lambda	.841	72.000	2062.37	.927	.983	.650

The assumption of non-correlation between dependent variables was tested using Pearson's r and reported in table 6.2.2.2 below. Where dependent variables are uncorrelated, univariate F-tests may be used to explore relationships between independent and dependent variables. Domain scores were indeed uncorrelated (table 6.2.2.2). Consequently, univariate F-tests were used to further explore the effect of trust and professional group on CGCQ factor scores.

Table 6.2.2.2: Correlations between dependent variables (Pearson's r)

	Quality improvement	Risk management	Absence of unjust blame	Working with colleagues	Training and development
Risk management	.039	-			
Absence of blame	.023	.045	-		
Working with colleagues	.052	.011	.034	-	
Training and development	.044	.045	.007	.055	-
Organisational learning	.044	.035	.012	.008	.024

All r's non-significant. Dependent variables are uncorrelated.

2. *If there is a significant multivariate effect which of the six CGCQ domains does trust or professional group affect? Which trusts or professional groups are significantly different to each other?*

(a) Trust

The full sample of 406 respondents included an additional 171 respondents who were not from any of the four sampled trusts but recruited from conference attendees at the Health Services Management Centre in order to boost sample size. Consequently, the additional 171 were dropped from the analysis; the effect of trust on dependent variables reported below is based on the remaining 235 cases. Given uncorrelated dependent variables, the effect of trust on CGCQ factor scores was explored further using univariate F-scores (table 6.2.2.3).

Table 6.2.2.3: Univariate F-scores of effect of trust on CGCQ factor scores

DV	Univariate F	d.f.	Sig.
Quality improvement	3.999**	3	.008
Risk management	2.208	3	.088
Absence of unjust blame	2.609	3	.052
Working with colleagues	1.384	3	.248
Training and development	5.337***	3	.001
Organisational learning	13.387***	3	.000

*** $p < .001$

** $p < .01$

Univariate analysis identified a significant effect of trust on scores for organisational learning ($F = 13.387$, d.f. = 3, $p < .001$), training and development ($F = 5.337$, d.f. = 3, $p = .001$) and quality improvement ($F = 3.999$, d.f. = 3, $p = .008$). Post-hoc comparisons using Bonferroni adjustment for multiple F tests were used to explore differences further and identify which trusts differed and in what way.

Table 6.2.2.4: Inter-trust means for organisational learning, training and development and quality improvement

Trust descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Organisational learning	Acute teaching	50	50.3127	12.51861	1.77040	46.7550	53.8705	17.27	74.95
	PCT	71	59.0531	11.19758	1.32891	56.4027	61.7036	32.85	91.62
	Community/MH/LD	68	54.6881	10.69893	1.29744	52.0984	57.2778	21.95	76.36
	Acute specialist	46	45.8899	12.68144	1.86978	42.1240	49.6558	.00	65.90
	Total	235	53.3538	12.54303	.81822	51.7417	54.9658	.00	91.62
Training and development	Acute teaching	50	38.7343	20.76071	2.93601	32.8342	44.6344	.00	77.96
	PCT	71	49.8880	16.11668	1.91270	46.0732	53.7027	5.09	90.33
	Community/MH/LD	68	49.2013	15.09746	1.83084	45.5469	52.8557	21.23	100.00
	Acute specialist	46	46.7379	13.12371	1.93499	42.8406	50.6352	13.69	77.57
	Total	235	46.6996	16.87604	1.10087	44.5307	48.8684	.00	100.00
Quality improvement	Acute teaching	50	42.6501	9.93509	1.40503	39.8266	45.4736	24.81	65.17
	PCT	71	48.2626	12.96461	1.53862	45.1939	51.3313	.00	68.41
	Community/MH/LD	68	44.8103	13.42590	1.62813	41.5606	48.0601	11.42	73.29
	Acute specialist	46	40.5814	13.06041	1.92565	36.7030	44.4599	12.16	67.01
	Total	235	44.5659	12.78409	.83394	42.9230	46.2089	.00	73.29

(i) Organisational learning

Both the acute teaching and acute specialist trusts scored lower than the PCT ($p < .001$); additionally, the specialist trust scored lower than the community/ MH/ LD trust ($p = .001$).

(ii) Training and development

The acute teaching trust scored lower than both the PCT ($p = .002$) and the community/ MH/ LD trust ($p = .004$).

(iii) Quality improvement

The acute specialist trust scored lower than the PCT ($p = .008$)

(b) Professional group

Given uncorrelated dependent variables, the effect of professional group on CGCQ factor scores was explored further using univariate F-scores (table 6.2.2.5).

Table 6.2.2.5: Univariate F-scores of effect of professional group on CGCQ factors

DV	Univariate F	d.f.	Sig.
Quality improvement	1.306	3	.272
Risk management	10.099***	3	.000
Absence of unjust blame	12.479***	3	.000
Working with colleagues	5.110**	3	.002
Training and development	1.850	3	.138
Organisational learning	1.722	3	.162

*** $p < .001$

** $p < .01$

Univariate analysis identified a significant effect of professional group on scores for absence of unjust blame ($F = 12.479$, d.f. = 3, $p < .001$), risk management ($F = 10.099$, d.f. = 3, $p < .001$), and working with colleagues ($F = 5.110$, d.f. = 3, $p = .002$). Post-hoc comparisons using Bonferroni adjustment for multiple F tests were used to explore differences further and identify which professional groups differed and in what way.

Table 6.2.2.6: inter-professional means for absence of blame, risk management and working with colleagues

Professional group descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Absence of blame	Medical	58	53.4075	15.33072	2.01302	49.3765	57.4385	23.60	80.27
	Nursing / PAM	151	59.2753	15.82393	1.28773	56.7309	61.8198	.00	99.33
	Managerial	129	66.0903	11.77103	1.03638	64.0396	68.1409	32.79	100.00
	Other	65	59.2592	11.65269	1.44534	56.3718	62.1466	34.27	91.36
	Total	403	60.6097	14.51132	.72286	59.1886	62.0307	.00	100.00
Risk management	Medical	58	45.8936	19.05077	2.50149	40.8845	50.9028	3.86	100.00
	Nursing / PAM	151	53.6760	14.08798	1.14646	51.4107	55.9413	20.07	85.30
	Managerial	129	59.1745	15.34315	1.35089	56.5016	61.8475	.00	90.37
	Other	65	53.3015	15.50038	1.92259	49.4607	57.1423	16.11	85.46
	Total	403	54.2556	16.02656	.79834	52.6862	55.8251	.00	100.00
Working with colleagues	Medical	58	53.6878	11.77892	1.54665	50.5907	56.7849	25.70	80.08
	Nursing / PAM	151	48.3593	12.61380	1.02650	46.3310	50.3875	14.54	100.00
	Managerial	129	46.4901	10.95008	.96410	44.5825	48.3977	21.78	78.17
	Other	65	47.4058	11.85884	1.47091	44.4674	50.3443	.00	66.79
	Total	403	48.3741	12.04048	.59978	47.1950	49.5531	.00	100.00

(n = 403) as 3 respondents did not record their professional group

(i) Absence of unjust blame

Managers scored higher than medics ($p < .001$), nurses / PAMs ($p < .001$) and others ($p = .008$); and nurses / PAMs scored higher than medics ($p = .040$). This suggests that relative to other professional groups, managers believe that there is less unfair blame.

(ii) Risk management

Managers scored higher than medics ($p < .001$) and nurses / PAMs ($p = .020$); and nurses / PAMs scored higher than medics ($p = .008$). This again suggests that managers are more optimistic than others are of routine risk management activity.

(iii) Working with colleagues

Medics scored higher than Managers did ($p = .001$), as did nurses / PAMs ($p = .020$) and others ($p = .021$). This suggests that medics perceive working relationships with others to be better than all other professional groups.

6.3 Discussion

Results provide evidence of differences in clinical governance climate between study sites, and consequently that the CGCQ is sensitive enough to identify inter-site differences. Yet, results showed that both organisational context *and* professional subgroup influence respondents' perceptions of clinical governance climate. Significant differences between study sites were found for the domains of quality improvement ($F = 3.999$, d.f. = 3, $p = .008$), training and development ($F = 5.337$, d.f. = 3, $p = .001$) and organisational learning ($F = 13.387$, d.f. = 3, $p = .000$). On absence of blame and risk management measures, respondents from the PCT site scored clinical governance climate more positively.

Similarly, there were significant differences between professional groups for the domains of risk management ($F = 10.099$, d.f. = 3, $p = .000$), absence of unjust blame ($F = 12.479$, d.f. = 3, $p = .000$) and working with colleagues ($F = 5.110$, d.f. = 3, $p = .002$). On each of these domains managers scored the clinical governance climate more positively than did doctors.

While differences between site scores on the Competing Values Questionnaire (CVQ) were small, statistically significant differences between sites were found. Scores for 'rational' were significantly lower for the PCT (20.70) than both the acute teaching (31.16) and acute specialist (28.78) hospitals. Thus PCT respondents perceived their organisation to be less characterised by planning and productivity than respondents from the acute sites. Scores for clan were significantly higher for the MH/LD/SC trust (25.24) than the acute teaching hospital (14.88). Thus MH/LD/SC trust respondents perceived their organisation to be more characterised by trust, tradition and member commitment.

Mannion and colleagues reported that hierarchical cultures were associated with better performance on 'star' ratings (2003a; 2003b). By extrapolation, the negative association between CGCQ domains and CVQ hierarchical cultures in the present study reported in chapter five above suggests that achievement in star ratings are associated with modes of management that will make the long term agenda of clinical governance more difficult to achieve. In short, those organisations feted for controlling waiting times may be exactly the ones least able to engage with the long-term quality improvement agenda. Rewarding reductions in waiting times may act as a perverse incentive further undermining the long-term agenda.

Results from the present study also showed significant differences between professional groups that were *consistent across trust*. There were significant differences between professional groups for the domains of risk management ($F = 10.099$, d.f. = 3, $p = .000$), absence of unjust blame ($F = 12.479$, d.f. = 3, $p = .000$) and working with colleagues ($F =$

5.110, d.f. = 3, $p = .002$). This suggests perceptions of clinical governance climate on these domains are mediated through respondents' professional discipline. Results showed that across the domains risk management and absence of unjust blame, managers were more positive than medical staff. Results for nurses / paramedics were less consistent across the domains. On absence of unjust blame and risk management, they scored significantly higher than medics but lower than managers did; on working with colleagues they scored significantly higher than medics did but were no different to managers.

The potential for inter-professional variation signals the importance of measuring perceptions of clinical governance culture across staff groups. Policy authorities intend that the clinical improvement agenda will be effected by local clinical governance arrangements that make trust boards directly accountable for health care quality assurance, clinical audit and clinical development. However, managers require the active participation of clinicians to implement these policy initiatives at the clinical level. Such active implementation is dependent upon how the professional groups interpret the policy imperatives, themselves dependent upon how clinicians conceptualise their clinical work. Those CGCQ domains for which site-based differences were found, such as quality improvement, training and development and organisational learning approaches, may be addressed corporately via HR strategy. However, there was evidence that professional groups experience the pressures associated with risk management, absence of unjust blame and working with colleagues differently.

Prior work by Degeling *et al* (2001; 2003) strongly suggests that medical, nursing and managerial staff have distinct profession-based conceptions of clinical work, which influence opinion concerning health care reform programmes. Distinctions include individualist versus systematised conceptions of clinical work and conceptions of the financial and accountability aspects of clinical work (i.e. between financial realism and transparent accountability and clinical purism and accountability to self, peers and patients). General managers tend to hold strongly systematised conceptions of clinical work, are supportive of financial realism and transparent accountability. In contrast, medical clinicians hold strongly individualist conceptions of clinical work, and are equivocal about financial realism and transparent accountability. While nurse clinicians shared general managers' systematised conceptions of clinical work, they were even less favourable towards transparent accountability than medical clinicians.

While Degeling *et. al.* (2003) confirm the attribution of tensions between doctors and managers to differing professional cultures, they also point to the potential contribution of nurse clinicians. Specifically, the multi-disciplinary team based systems favoured by nursing staff groups, such as clinical pathways for diagnostic groups, could be used as the basis of re-establishing 'responsible autonomy' as the organising principle of clinical work:

"... integrated care pathways specify the agreed sequence of diagnostic and therapeutic processes and incorporate the views of (medical, nursing, and allied health) clinicians and managers, which (in the light of the available evidence, stated resource constraints and experience of patients) are essential for achieving desired outcomes for specified clinical conditions. They therefore provide a tool for empowering clinicians to strike a balance between the clinical and resource dimensions of care and between the requirements of clinical autonomy and transparent accountability." (Degeling *et. al.* 2003: p. 651)

Failure to do so however could leave policy authorities with no alternative but to continue with a proliferation of accountability systems and regulatory structures, leading to a further cycle of clinician resistance (Degeling *et. al.* 2001). Results of the present study suggest that nursing and paramedical staff have the potential to act as bridging agents between managers and medics across the important domains of risk management, absence of unjust blame and working with colleagues.

It is often stated that continuous quality improvement requires the generation of “real commitment and enthusiasm” for quality improvement in all staff “transcending departmental and disciplinary boundaries” (Pollitt 1996: 105). Results suggest that in the study sites, professional differences have not converged. Yet, there is a need to avoid the naivety of a simple integrationist perspective (Martin 2002) and to take care not to overstate the need for newness: to honour and retain what is currently valuable.

“In attempting to help organizations develop ‘a strong unitary culture’, [management] consultants often fail to look at the richness of the culture as it already exists with its multiple and complex meaning-making processes”
(Hawkins 1997:425)

The culture change literature reviewed in chapter four urged caution on the part of executives charged with cultural change. The worry is that attempts to superimpose a set of new espoused values on subordinates that are discordant with the latter’s experience of reality will lead subordinates act out the surface signals of the ‘new culture’ cynically and

without internalisation. This may confirm senior management that the culture has taken- but in reality all that has happened is a resigned compliance to the surface messages of the new culture underpinned by 'business as usual':

"Without negative feedback senior management is reinforced in the delivery of a message that may be increasingly at odds with employees' perceptions of organizational reality, and engender an increasing cynicism. In such circumstances the paradox will exist that as senior management becomes increasingly convinced that cultural change has taken place, other organizational members may become increasingly cynical and resistant to such attempts to change behaviour."
(Legge 1994: p. 125)

Legge (1994) and Anthony (1990) question both the ability and desirability of managerial leadership to transform clinical subcultures. Clinical cultures are not imposed by the institution- they are based upon widely shared professional values. Thus clinicians traditionally view institutions as instrumental to clinical practice - *the institution is secondary to the practice*. This may partially explain the resilience of such professional cultures to crude attempts at 'cultural makeover'. There is also an attendant ethical dilemma involved in the development of a strong culture through transformational managerial leadership in professional bureaucracies (Legge 1994). Such leaders initially represent a sub-group engaged in battle with a strong (but perceived ineffective) dominant culture. In the case of the NHS, managers may be perceived to represent the espoused management values of the primacy of the efficient allocation of resources, with professional clinicians arguing the primacy of the delivery of best patient care as defined by

clinicians. Given the tradition of moral concern within clinical practice, the legitimacy of such cultural transformation is at least questionable:

"... we have the paradox that a small subgroup (senior managers) is being called on to create a strong subculture to confront and overwhelm a more widely shared strong culture, one that is rooted in moral rather than instrumental values. Questions inevitably must be raised about the ethics of such initiatives..."
(Legge 1994: p. 125)

Naive cultural change initiatives will be regarded as corporate culture rather than organisational culture; a transparent management concoction, imposed on staff. Staff are not passive consumers of corporate culture: indeed, staff will actively engage with it, either to accept or reject (Wright 1994). Rhetoric aside, the clinical governance agenda calls on clinicians to balance their clinical autonomy with accountability to management, recognise the resource dimensions of clinical decisions, subscribe to team-based approaches to clinical work and support increased systematisation of clinical work (Degeling *et. al.* 2003).

There is a fundamental paradox between calls for a common set of values between clinicians and managers, and a belief that doctors and managers may legitimately have different concerns. Indeed, people may be best served by the tension between clinician's concern for the needs of the individual patient irrespective of the needs of others or of budgets, and the managers concern for the delivery of policy and the aggregated needs of populations (Edwards *et. al.* 2003). However, it may also be possible to dialectically resolve the conflicts between managers and medical staff by drawing in the broader clinical team to establish 'responsible autonomy' (Degeling 2003). Failure to dialectically produce

responsible autonomy, and to rely on bureaucracy to police the system, is to invite further resistance (Davies *et. al.* 2001), rendering the attempt at cultural management self-defeating.

Conclusions

This chapter reported results of an application of the CGCQ schedule in a sample of 235 respondents from four health care trusts and an additional sample of 171 respondents for a profession-based analysis. Substantive results revealed statistically significant differences between both professional groups and trust sites on CGCQ domain scores. Consequently, perceived clinical governance climate is mediated through professional status in the domains of absence of unjust blame, risk management and working with colleagues. Differences between professional perspectives indicate the existence of multiple perceived 'clinical governance climates' within organisations, and those charged with implementing clinical governance need to be aware of such differences of perception when designing interventions.

CHAPTER 7

MEASURES OF CLINICAL GOVERNANCE II:

The Organisational Progress in Clinical Governance (OPCG) schedule

Introduction

Options for measuring progress in clinical governance were reviewed in chapter 4 above, and a case made for the development of two new and complementary measures, operating at different units of analysis, to support internal organisational development. The first, a clinical governance climate questionnaire (CGCQ), enables exploration of individual staff experience; the second, an organisational progress in clinical governance (OPCG) competency framework, facilitates an examination of perceptions of progress at board and directorate level. This chapter opens with a précis of the research programme used to derive the OPCG schedule and assess its reliability and validity, before considering each stage in detail.

7.1 Developing a competency framework

The OPCG schedule is designed to provide organisations with information on perceived progress on a number of dimensions related to good clinical governance, at board and directorate levels.

7.1.1 An overview of OPCG development methods

The importance of identifying tasks and levels of attainment against these tasks, within competency frameworks cannot be overstated. An empirically derived competency framework requires good practice to be operationalised in terms of tasks for completion; the generation of statements denoting degrees of competence in the tasks; and validation of the level of competency represented by each statement. This section provides an overview of the research programme used in each of these stages of development of the OPCG competency framework. Subsequent sections provide detailed information on each component in sequence, including an assessment of the methods chosen and the methodological decisions taken in the context of the present study, as well as substantive results. Table 7.1.1.1 below contains an overview of the programme, for purposes of clarity.

(a) Operationalising good practice: generating the tasks to be completed

As clinical governance is an abstract construct, the first difficulty encountered in developing a competency framework is that of operationalising good practice. Possible research-based options include the review of related primary literature(s), or primary research with interest groups using qualitative methods. This study combined both approaches, by initially reviewing clinical governance policy documents in order to generate tasks defining good practice, then using qualitative methods with expert samples to operationalise levels of competence within the identified tasks.

Table 7.1.1.1: OPCG competency framework research programme

Stage	Aim	Process
(a) Identify the tasks	To operationalise 'good practice' in clinical governance	1. Review clinical governance policy documents 2. Identify major development areas and tasks
(b) Generate competence statements under each task	(i) To operationalise competency levels	1. Recruit an expert panel 2. Panel generate Q-sample statements of functioning at different levels of competency for each task identified in (a) above
	(ii) To empirically derive levels of competency for statements in the framework	1. Recruit an expert panel 2. Panel Q-sort the Q-sample generated in (b) above to assign a level of competency to each statement, using a forced-free response grid 3. Examine data for consensus opinion on competency levels
(c) Reliability and validity testing	To empirically assess the factor structure, internal consistency, and validity of the OPCG	1. Postal questionnaire (n = 1916) 2. Forced- factor factor analysis to assess the internal consistency and underlying factor structure 3. Alpha-coefficients for internal consistency 4. Construct validation of the competency levels assigned to items

The vast and multi-dimensional nature of literatures related to clinical governance necessitated a decision on where to draw the boundaries for the literature review. For the purpose of operationalising good practice, review boundaries were drawn tightly around policy documents, excluding the literature detailing conceptual disputes previously explored in chapter one. Given the broad sweep of policy pronouncements, this decision did not result in a narrow subset of topics. Rather, it focused attention clearly on broad policy aspirations such as managing risks, improving quality and leadership and collaboration, which gives the resultant framework good face validity with service staff, further encouraging its use. Once good practice was operationalised in terms of policy intentions, the activities implied by these intentions were themselves operationalised by generating a series of concrete tasks. For example, the policy intention of managing clinical risk implied the task of developing and using organisation-wide risk management systems, among others. Such tasks form the basis of the competency framework, denoting good practice in clinical governance in terms of the activities required. Section 7.1.2 below provides details of the review of government guidance and the tasks subsequently derived.

(b) Operationalising competencies

Generating competence statements for each task

Following the generation of tasks denoting good practice in clinical governance, the next stage required development of statements representing degrees of competence in the tasks. This was approached empirically via qualitative exploration of the tasks with an expert sample of participants. After considering a range of qualitative techniques such as Nominal Group Technique (NGT), Q-sort and Delphi, the preferred approach was a Q-sort study, in

which an expert group generated statements for the Q-sample, later Q-sorted by a separate expert panel. The principal advantage of the method in the context of the present study is its ability to derive and sort statements of achievement into levels of competency.

After the decision was taken to use Q-sorting as the method for assigning item descriptions of tasks to levels of competency, the first difficulty was to generate items describing achievement against the tasks at different levels of competency. An expert group was convened, consisting of participants reflecting the range of intended users of the final measure, and including chief executives, executive and non-executive board members, clinical directors and operational managers. Participants worked in two groups generating statements of competence against the tasks for the planned Q-sort. In response to task definitions, participants generated statements related to functioning at different levels of competence for each task. Collectively, these statements formed the Q-sample.

Assigning levels of competency to the statements

Q-sorts require an expert panel of respondents to sort a series of statements on a topic of interest into a pre-defined distribution. Each respondent receives a set of statements printed on separate cards, to sort into categories. In this application, respondents sorted the Q-sample generated by the initial expert panel into categories, based on the level of competence in clinical governance represented by each statement. A forced-free allocation was used, in which participants placed an equal number of statements under six headings, numbered from one (low competency) to six (high competency). This represents the normative rating of statements relative to each other. The heading under which each

participant recorded each statement was recorded, and where there was broad agreement between participants on the level of competency denoted by a statement, it was retained. The final framework consisted of 54 of the original 144 statements.

(c) Reliability and validity testing

A forced five-factor exploratory principal factor analysis was used to examine the internal consistency of OPCG items within five anticipated domains of improving quality. The domains (managing risks, improving performance, corporate accountability and leadership and collaboration) were based on groupings of tasks derived from the literature review. The five empirical factors matched the anticipated domains, and individual items loaded on their expected domain. Residuals were low, which suggested that the five-factor solution was appropriate and no further unextracted factors remained. All five domains had coefficient alpha values between 0.78-0.92, which offers further support to the internal consistency of the five OPCG domains. The content validity of the OPCG is attested by the use of literature reviews to set the boundaries for further development and qualitative research with nominal expert groups to generate tasks and levels of competency. The ability of the OPCG to discriminate between achievement on aggregations of items previously designated as indicative of 'low' difficulty from those designated 'medium' and 'high' difficulty lends qualified support to its construct validity.

7.1.2 Operationalising good practice: generating the tasks

As outlined briefly above, the vast and multi-dimensional nature of literatures related to clinical governance necessitated a decision on where to draw the boundaries of a literature

review to inform OPCG development. The broader literature, as reviewed in chapter one, provides insight into disputes surrounding the development and implementation of clinical governance. However, it is less useful for the purpose of operationalising good practice. The need for clarity on the major areas of importance informed the decision to return *ad fontes* and review the three government policy documents specifically relating to clinical governance. These three consist of the 1997 White Paper *The new NHS: modern, dependable*; the discussion document *A first class service: Quality in the new NHS* (1998); and the guidance document *Clinical governance: Quality in the new NHS* (1999).

The White Paper provides a high level statement of intent for major changes in the delivery of NHS health care. While the document briefly outlines new roles and responsibilities of trusts within the context of the new NHS in chapter six, there is little detail of what might constitute good practice in clinical governance. Rather, it identifies clinical governance as the mechanism by which health care quality is to be placed at the core of corporate and professional responsibilities. The former is to be achieved via a statutory duty for quality for CEOs and local structural arrangements, and the latter via strengthening existing systems of self regulation to tackle poor performance and manage clinical risk (DoH 1997). A little more detail is forthcoming in section 7.12 of the report, which lists a limited number of attributes of a 'quality organisation' with the system. Attributes include quality improvement processes; leadership skills; EBP; dissemination of learning; risk, adverse event and complaints management; management of poor performance; professional development; and collection of quality clinical data. While these provide some indication of the required elements, it remains impressionistic.

The discussion and guidance documents provide much more detail. They list the main components of clinical governance for trust under four headings: accountability; quality improvement; risk management; and identifying (and remedying) poor performance (Table 7.1.2.1). The later guidance document, *Clinical governance: quality in the new NHS*, additionally reports a number of long-term benefits expected to result from clinical governance implementation in an annex to the report. These are absent from the earlier discussion document, and their presence in the later guidance is evidence of the desire of those responding to the discussion document to encourage a continuous quality improvement focus. These benefits concern the involvement of users and carers, dissemination of good practice across and between organisations, and explicitly the proactive support and development of clinical staff (Table 7.1.2.2).

In order to construct a competency framework, these government document components were re-expressed as a series of 20 concrete tasks, organised under five conceptual headings: improving quality; managing risks; improving performance; corporate accountability; and leadership & collaboration (Table 7.1.3.3). These 20 tasks formed the basis of further development of the OPCG competency framework, consisting of a Q-Sort study in which expert groups generated and validated statements of levels of achievement against the tasks.

Table 7.1.2.1: OPCG Task areas mapped on to 'Main components of clinical governance for NHS Trusts' from discussion and guidance documents, DoH 1998/9

Government document component	OPCG task
1. Clear lines of responsibility and accountability for the overall quality of clinical care	
• The NHS Trust Chief Executive carries ultimate responsibility for assuring the quality of services provided by the trust	13
• A designated senior clinician responsible for ensuring that systems of clinical governance are in place and monitoring their continued effectiveness	13
• Formal arrangements for NHS Trust Boards to discharge their responsibilities for clinical quality, through a clinical governance committee	14
• Regular reports to NHS Trust Boards on the quality of clinical care given the same importance as monthly financial reports	13
• An annual report on clinical governance	13
2. A comprehensive programme of quality improvement activities	3
• Full participation of all hospital doctors in audit programmes, including speciality and sub-speciality national audit programmes endorsed by the Commission for Health Improvement	4
• Full participation in the current four National Confidential Inquiries	4
• Evidence-based practice is supported and applied routinely in everyday practice	2
• Ensuring the clinical standards of National Service Frameworks and NICE recommendations are implemented	16
• Workforce planning and development is fully integrated within the organisation's service planning	15
• Continuing Professional Development: programmes aimed at meeting the development needs of individual health professionals and the service needs of the organisation are in place and supported locally	11
• Appropriate safeguards to govern access to and storage of confidential patient information as recommended in the Caldicott Report on the Review of Patient-Identifiable Information	5
• Effective monitoring of clinical care with high quality systems for clinical record keeping and the collection of relevant information	5
• Processes for assuring the quality of clinical care are in place and integrated with the quality programme for the organisation as a whole	15
• Participation in well-designed, relevant R&D activity is encouraged	5

Government document component	OPCG task
and supported as something which can contribute to the development of an 'evaluation culture'	
3. Clear policies aimed at managing risks	
• Controls assurance which promotes self-assessment to identify and manage risks	6
• Clinical risk systematically assessed with programmes in place to reduce risk	6 & 9
4. Procedures for all professional groups to identify and remedy poor performance	
• Critical incident reporting ensures that adverse events are identified, openly investigated lessons learned and promptly applied	7 & 9
• Complaints procedures, accessible to patients and their families and fair to staff. Lessons are learned and recurrence of similar problems avoided	8 & 9
• Professional performance procedures which take effect at an early stage before patients are harmed and which help the individual to improve their performance whenever possible, are in place and understood by all staff	11
• Staff supported in their duty to report any concerns about colleagues' professional conduct and performance, with clear statements from the Board on what is expected of all staff. Clear procedures for reporting concerns so that early action can be taken to support the individual to remedy the situation	10

Table 7.1.2.2: Benefits arising from clinical governance- the vision for the next five years (after *Clinical governance: quality in the new NHS*)

Benefits	OPCG task
5. Involving users and carers	18
• An organisation-wide strategy for involving patient, users, carers and the public, including strategic plans for communicating with them • User representatives on clinical governance committee / groups • Use of involvement methodologies e.g. patient panels, focus groups • Training and education for all individuals on effective patient, user, carer and public involvement	
6. Sharing of good practice	1 & 19 & 20
• Evidence that individuals and organisations are actively learning from others, for example by actively seeking out and making use of examples of good practice and of the ways in which particular issues have been tackled elsewhere	
7. Detecting and dealing with poor performance & adverse events	
• Clear mechanisms for the identification and management of poorly performing clinicians; routes for other clinicians to voice their concerns about performance of their colleagues (taking account of new national policies and procedures)	11
• Poor performance procedures aim to identify practice as it begins to slip, enabling sustained improvements in the vast majority of cases without risk to the quality of patient care	12

Table 7.1.2.3: The 20 OPGC tasks, derived from discussion and guidance documents

- | |
|--|
| <p>a) Improving quality</p> <ol style="list-style-type: none">1. Benchmarking practice2. Using research evidence3. Discussing clinical issues4. Auditing clinical practice5. Using performance indicators <p>b) Managing risks</p> <ol style="list-style-type: none">6. Using risk management systems7. Using adverse events systems8. Using complaints systems9. Learning from complaints, adverse events and near misses10. Raising issues in confidence <p>C) Improving performance</p> <ol style="list-style-type: none">11. Professional development for staff12. Appraising staff <p>D) Corporate accountability</p> <ol style="list-style-type: none">13. Accountability for clinical quality14. Organising clinical governance committee structures <p>E) Leadership and collaboration</p> <ol style="list-style-type: none">15. Linking development activity to the clinical governance agenda16. Using National Service Frameworks (NSFs) & guidelines17. Leadership and team building18. Involving service users in service development19. Collaborating with local health and social care partners20. Developing organisational learning |
|--|

7.1.3 Operationalising levels of achievement

This section offers a review of the options considered for deriving levels of achievement within the competency framework including Delphi, nominal-group and Q-sorts. After outlining the reasons for selecting a Q-sort approach, the section provides a critical

introduction to Q-methodology and Q-sorting, and explores its specific application within the present study.

7.1.3.1 Choice of Q-sort in the present study

A method was required to generate a series of statements representing achievement against each task, and assign a level of competency in the tasks between 'low' and 'high' represented by each statement. The essential consideration is the need to rank the level of achievement represented by a statement, relative to all other statements. A scoring system in which the scoring of each item is related to and dependent on the other items is *ipsitive* (Karim 2001). This is particularly important in the current application, which requires the designation of items into different levels of competency. We need to be confident that all of the items seen as evidence of achievement at 'low', 'medium' or 'high' levels are comparable. Unless items are norm referenced against each other in this way, we could not know whether items designated 'low' in one task represented comparable achievement to other items designated 'low' *against different tasks*. The sorting of all items relative to each other meets the above requirement and increases confidence in the equivalence of the designations of 'low', 'medium' and 'high' across all items. In subsequent data collection, this equivalence allows the meaningful aggregation of scores of progress across all 'low', 'medium' and 'high' items into a single summary score for each respondent at each of the three levels.

Delphi, Nominal Group Technique (NGT) and Q-sorts all offered the potential to combine statement generation and validation of competency level in a single method. Yet, both Delphi and NGT studies faced particular problems. The present application required each

statement to be norm referenced against each other; to be sure that an item designated 'low' in a task represented comparable achievement to other items designated 'low' against *different tasks*. The sorting of all items relative to each other increases confidence in the equivalence of the designations 'low', 'medium' and 'high' across all items. While both NGT and Q-sorts explicitly involve normative rating, it is difficult to apply in Delphi scoring, which typically requires individual rating of items against a Lickert scale (Mullen 2000). Together with the propensity for large attrition between voting rounds in Delphi studies, this made a multi-round Delphi study unattractive. While NGT is efficient in generating and ranking items normatively in a single meeting, the potentially very large number of items for rating in this application made running such a group impractical. On balance, Q-sort represented the best alternative, with its normative scoring of items and ability to cope with large numbers of items simultaneously. The separation of item generation and rating into different rounds increased the likelihood of successful completion. While often used within the context of a broader 'Q-methodology' approach, Q-sort is simply a data collection technique.

Definitions and uses of Q-methodology

Q-methodology or Q-method is a pattern analytical approach, representing an attempt to analyse subjectivity in a structured, statistical form (Brown 1996). The basic distinctiveness of Q-method is that it is interested in establishing patterns *within and across individuals*, rather than across individual traits (such as gender or age) as in standard survey analysis. In essence, it attempts to elicit the variety of expressed opinion around an issue or topic. It operates on the assumption of 'finite diversity' i.e. that there will be a limit to both

the number of ordered patterns shared across individuals and the diversity of accounts, and reveals them in a structured manner.

Originally devised in the 1930s by Stephenson (1953) and used within the discipline of psychology, the method has had considerable influence in the study of political science in the USA since the 1980s (Brown 1980). More recently in the UK the method has been embraced and developed as an approach particularly suited to a social constructionist research paradigm, especially where concerned with the expression of a diverse range of views (Stainton Rogers 1991).

Stephenson's original exposition of Q-method describes its ability to facilitate the expression of subjectivity on a given topic, a sense of life as lived from the perspective of the person living it. This is achieved via the medium of sorting a given number of items under specific conditions of instruction. Thus for Stephenson, Q-method allows the analyst direct access to human expression, rather than seeking to tap essences as in conventional psychometrics. This reading of Q-method is compatible with a realist ontology i.e. the belief in a pre-existent individual viewpoint, to which the technique gains access. Recent healthcare applications include exploring the needs of support staff in emergency medicine (Chinnis *et. al.* 2001); identifying stakeholder perceptions in strategic planning (Popovich and Popovich 2000); and exploring variation in adoption of information technologies (Valenta and Wigger 1997).

Social constructionist applications move beyond Stephenson's position, and deny the notion of an underlying individual viewpoint. This reading regards Q-method results as culturally available discourse on the topic in question, in which 'discourse' means a way of seeing and talking about a topic. The essence of discourses is that they are both individual and subjective and represent the way an individual, at a particular time and circumstances, forms conceptions about the world (Stainton Rogers 1995). In this social constructionist reading, Q-method is concerned with uncovering patterns of shared (social) discourses, which it does by correlating individual responses to extract latent shared discourses using Q-factor analysis (Barry and Proops 1999). Emergent factors are examples of the extent to which participants draw on the available discourses, expressed through the Q-sort. Such discourses are fluid, and it is to be expected that the expressions will vary over time and context.

Within the UK, social constructionist applications have proliferated in gender studies and health care research. Feminist applications include subjectivity and diversity in accounts of lesbianism (Kitzinger 1986); women's experiences with and attitudes towards pornography (Senn 1993); male identity (Horwood 2000) and feminism itself (Snelling 1999). Health care examples include experiences of Neuroleptic medication (Day *et. al.* 1996); Irritable Bowel Syndrome (Stenner *et. al.* 2000); managing risk in mental health services (McKeown *et. al.* 1999) and nurses' conceptions of industrial relations (McKeown *et. al.* 1999).

Q-sort

While Q-method as developed by social constructionists takes a distinctive ontological stance, its instrumental basis, the Q-sort, is a *data collection technique* which has been widely applied by researchers from different research traditions, including standard psychometrics (Block 1978; Prasad 2001). It is essentially a technique used to measure the position of concepts or individuals relative to each other, conventionally involving the ordering of a set of statements from agree strongly to disagree strongly.

The flexibility of Q-sort technique, reflected in the breadth of applications across disciplines outlined above, is further highlighted by the variety of approaches to data analysis used in applications. Brouwer (1999) provides a helpful typology, noting its use in intrapersonal (single case) and interpersonal (population based) studies. The former intrapersonal applications make use of multiple sorts of the same items by the same individual, using variations of the same basic conditions of instruction for each sort. A classic application is Rogers's (1954) assessment of real vs. ideal self-descriptions. An individual sorted items based on their accuracy as descriptions of 'my real self', under headings ranging from 'agree strongly' to 'disagree strongly', and then sorted the items again based on their accuracy as descriptions of 'my ideal self'. Following collection, Rogers (1954) examined the correlation between the two sorts. He identified a strong correlation between the two sets of results with 'maturity', in the sense of self-acceptance. In contrast, interpersonal studies make use of a single sort of the items by multiple individuals under the same conditions of instruction in order to establish response patterns across individuals. Stenner *et. al.* (2000) provide a good example in their study of the range of understandings

about the nature and causality of their illness offered by people with Irritable Bowel Syndrome (IBS). Sixty respondents Q-sorted 58 statements about living with the condition under headings ranging from 'most agree' to 'most disagree'. This study used Q-factor analysis to identify seven clear and distinctive discourses used by respondents in association with the lived experience of IBS.

7.1.3.2 Q-sort study methodology

The brief review of government guidance documents in section 7.1.2 above identified a number of tasks considered important in clinical governance. A Q-sort study was chosen to develop these task areas into an organisational competency framework, by generating statements specifying different levels of achievement against each of the 20 tasks. The two research questions were clear:

1. What set of statements denote achievement of the tasks required for good clinical governance?
2. What level of attainment against a task does each statement represent?

Notwithstanding ontological disputes over the status of the results of Q-method, there is a consensus over the basic procedures for conducting Q-sorts (Stephenson 1954; Brown 1980; Stainton Rogers 1991; McKeown and Thomas 1988). These include deriving items for the Q-sample; selection of a participant sample; the setting of conditions of instruction; and specification of forced allocation. This section explores each element in sequence in the context of the present study,

Deriving the Q sample

Also known as a Q-set or concourse, the Q-sample consists of a set of stimuli, each printed on a separate card. Typically comprising verbal statements, Q-samples may also be composed of pictures, recordings, or other stimulus material that is amenable to appraisal by the participant sample. The contents of the items are derived from interviews, literature reviews or other sources of information on the topic in question, and it is important that they cover the phenomenon of interest broadly, and contain the fullest range of viewpoints. In practice, this consideration is held in tension by the need to keep the number of items within manageable bounds. Large numbers of items pose practical problems for participants, as they are required to sort items into categories reflecting their score relative to all other items. This increases in difficulty relative to the number of items in the Q-sample. Typically, the Q-sample contains between 30 and 80 items, although applications of up to 100 are reported (McKeown and Thomas 1988).

In the present study, a Q-sample was required that consisted of statements representing degrees of competence in the identified tasks. In order to generate these statements for Q-sorting, an expert panel was convened. In this case, it was important that the assembled Q-set included the perspectives of a diverse range of professionals involved in the clinical governance agenda. To this end, the expert panel consisted of 12 members with specialist knowledge of clinical governance. Participants reflected the range of potential users of the final competency framework, including executive and non-executive board members,

clinical directors and operational managers with a responsibility for clinical governance in NHS Trusts (Table 7.1.3.2.1).

Table 7.1.3.2.1: Q-set generation study participants

Chief executives	1
Other executive directors	2
Non-executive directors	2
Clinician	3
Operational managers	4
	(n) = 12

Participants were purposively sampled, based on their known interest in the topic area and to ensure representation of managerial and clinical groups. On acceptance, participants were invited to a half-day seminar, at which they received further information on the purpose of the event and its place within the broader research programme. To make the task manageable, participants were sub-divided into two groups. Participant groups each received 10 sheets of paper. Each sheet listed one of the 20 tasks identified as required for the operation of good clinical governance based on the review of government guidance documents described above in section 7.1.2. On each sheet, respondents were encouraged to generate a series of statements reflecting achievement of the specified task at different levels of competency. Across the 20 identified task areas, participants' generated 144 statements (appendix 11).

Selecting the participant sample

The aim of recruitment for Q-methodological research is not to ensure a sample representative of a broader population; rather, it is to ensure that a wide range of relevant

viewpoints are represented (Snelling 1999). Thus in her study of women's perspectives on feminism, Snelling (1999) selected participants purposefully in order to ensure that a range of diverse perceptions were represented. Similarly, in their study of risk management and quality training, McKeown *et. al.* involved a diverse multi-professional, multi-agency mix of participants. The present study recruited operational managers with responsibility for promoting the NHS clinical governance agenda because of their expert knowledge on clinical governance. Given their central role in implementing clinical governance in NHS trusts, operational clinical governance managers were identified as the group with the most expert knowledge of the difficulties involved in achieving the competency statements generated by the Q-set expert group. Given the importance of diversity of opinion, it was important that participants were drawn from a broad range of different types of NHS trust. The expert panel consisted of eight members: three from acute trusts; two from mental health / learning disabilities trusts; two from primary care organisations; and one from an ambulance trust. Panel members were sampled purposively from managers known by the researcher to be particularly interested in the topic area due to previous contact on clinical governance related work.

Running the Q-sort group

Following an initial recruitment mailing, the expert panel was convened at the Health Services Management Centre at the University of Birmingham for a morning seminar. Participants received a clear brief for the morning, and detailed conditions of instruction (McKeown and Thompson 1988). As I wanted to explore the views of an expert group on the level of competence in tasks reflected by the items, the study utilised an interpersonal

design and a single, simple sorting request. Participants were briefed on the aim of the exercise and its importance within the overall research programme, in which the importance of their own judgements on the level of difficulty represented by each of the statements was emphasised. Each participant received a 'participant pack'. This consisted of two sealed envelopes containing the statements generated by the earlier expert group on cards, and two scorecards for recording the ID number of each statement card under one of six headings for each of two Q-sorts. Detailed instructions of the procedures used are contained in appendix 10.

Setting conditions of instruction

When sorting items in the Q-sample relative to each other, participants need clear instruction on the basis of comparison. In Q-sort terminology, these are termed conditions of instruction. They most often relate to a simple request, for example to sort items under ten headings running from 'those I most agree with (+5)' to 'those I most disagree with (-5)'. A variant form is the application of ranking continuum applied to differing theoretical constructs. For example, participants may be asked to sort the Q-sample items according to 'what is most like a liberal point of view (+5)' to 'what is most unlike a liberal point of view (-5)'. Whatever the chosen conditions of instruction, they need to be related to the principal research question, and unambiguous to participants. In the present study, participants sorted items into six categories on a continuum from those easiest to achieve and representing lowest competence in clinical governance (1), to those most difficult to achieve and representing the highest level of competence in clinical governance (6).

Specification of forced allocation

Usually, a Q-sort utilises a forced-free allocation, in which items are rank ordered along a continuum defined by the conditions of instruction, *allocating a prescribed number of items under each point on the continuum*. The allocation is forced to the extent that the number of items under each point is prescribed; and free to the extent that subject to the required distribution any given item may be allocated to any given point. A pyramidal structure approximating to a normal distribution is typical in forced-free allocations, to ensure that participants assign proportionately more items to the middle of the distribution. Such a normal distribution is not however required by the statistical techniques employed in factoring the data, and its use is primarily to encourage careful consideration in ranking (McKeown and Thomas 1988). In the present study, the Q-sort was required to identify roughly equal numbers of items of low, medium and high competency. Given this requirement, the allocation distribution chosen was a prescription of equal number of items across six points in a continuum from 'easiest to achieve' to 'most difficult to achieve'. The former represented operation at low levels of competence, and the latter represented operation at high levels of competence. In analysis, items assigned under points one and two in the continuum were equated with evidence of operation at 'low competency'; three and four with 'medium competency'; and five and six with 'high competency'.

The expert panel engaged in concourse development identified 144 items for the Q-set, which is twice the typical size of reported examples. It required respondents to make

comparisons between each item and the other 143 items, a task considered too difficult. Rather than arbitrarily reducing the number of items in the Q-set, it was divided into two equal Q-sets, each consisting of 72 items. Participants conducted separate Q-sorts on each. Following completion of both of the Q-sorts, participants had the opportunity to reassign items across the two Q-sorts if they wished. Thus if a participant believed that an item assigned under heading '5' of Q-sort one was actually more difficult to achieve than an item under heading '6' of Q-sort two, these items could be swapped. This allowed all 144 items to be sorted against each other, and made the task as manageable as possible.

7.1.3.3 Results

Participants sorted two Q-sets of 72 statements into six categories, from one (low competence) to six (high competence), and recorded the category that they had allocated each statement to. Where there was substantial agreement between participants on the category of a statement, it was retained for the final framework. Appendix 11 shows the number of times participants assigned each statement to each category, whether it was retained for the final framework and the category (level of competence) that the statement represents. Of 144 statements, 54 (37.5%) were retained.

Where statements were vague, scoring disagreements ensued. This was particularly noticeable in 6.1, 7.1, and 8.1 where participants were asked to rate statements about the robustness of risk management, adverse events and complaints systems. The difficulty here is that 'robustness' can be subjectively interpreted hence the wide spread of scores between 1(low) and 6 (high) between participants on these items.

Scoring rules for retention included a score of at least 4 in any aggregated segment (1,2,3). Additionally, there should be neither any more than 2 in any other aggregated segment or more than 2 other single segments (1,2,3,4,5,6). This was to identify statements where there was a wide spread of participants scores, as well as those where there were two distinct views on attainment difficulty. Additionally, consideration was taken of the number of statements from any one task retained, and the overall number of low, medium and high statements in the framework. Thus a delicate balance was required between agreement, coverage and balance of items.

It should be noted that some areas, such as involving service users and collaboration, all consist of 'high' competency items. Others, such as risk management, have a range of competency items low to high because they contain both simple 'collection' and complex 'organisational change' components. Others still, such as audit, are seen as relatively simple and all of their related statements are considered 'low' competency. Given that some areas were considered more difficult than others were there was no attempt to construct a 'matrix' of low, medium and high competency items for each area. However, given the Q-sort we can be sure that the level of competency assigned reflects differences in perception of achieving the item compared to all others across the framework.

Table 7.1.3.3.1: Q-sort items retained

1. Teams work together across boundaries to address quality issues
2. Staff identify best practice through benchmarking against other providers
3. Where appropriate, staff modify their care processes to reflect the best practice of others
4. Research evidence is regularly reviewed and discussed
5. Training in evidence based practice and critical appraisal is available to staff
6. Research Information is used consistently to inform our approach to quality improvement
7. Clinical issues are raised for discussion
8. Discussions on clinical issues are not dominated by any single profession
9. Staff participate in audit activity
10. Training in audit is available
11. Topics for audit are selected according to their potential impact on care quality
12. Risk management data is regularly reviewed and discussed
13. Following identification of a problem from risk data, clinical quality is improved
14. Clear action plans are developed in response to identified clinical risks
15. Staff receive training in using the risk management system
16. Data from the adverse event reporting system is regularly reviewed and discussed
17. Clear action plans are developed in response to adverse incidents
18. Following identification of a problem from adverse events data, clinical quality is improved
19. Staff are trained to use adverse event systems
20. Complaints are collated
21. Following identification of a problem from complaints data, clinical quality is improved
22. There is a 'no-blame' culture around reporting adverse events and near misses
23. Staff can raise clinical concerns about their colleagues in confidence
24. There is good access to agreed clinical performance indicators
25. Clinical indicators are used to reflect on, review and integrate services
26. Staff have development plans which identify training and development opportunities
27. Training identified in staff development plans matches individual needs to organisational needs
28. New skills gained through development activity are used in clinical settings
29. There is an annual staff appraisal process for most staff
30. An agreed work and development programme is used as the basis of staff appraisal
31. Staff appraisal is used as an opportunity to reflect on progress and plan future development
32. There is an executive director with responsibility for developing the clinical governance agenda
33. Quality improvement activity is planned and is included in a service development plan
34. Clinical areas have a nominated lead for clinical governance
35. There is a formal clinical governance committee, reporting to the board
36. There are local arrangements to collate information for the clinical governance committee
37. Organisation-wide clinical governance systems are underpinned by systems in clinical areas
38. Decisions about service developments or cutbacks are made on clear criteria
39. Local and national priorities from NSFs and HImPs are used to priorities service development
40. NSF implementation is integrated with business planning and quality improvement programmes
41. External guidelines are critically appraised before local adoption or development
42. Clinical protocols are shared with staff who work outside this organisation
43. The organisation shares a common vision for clinical governance
44. Leadership skills are identified and developed through leadership training programmes
45. Clinical teams receive performance feedback
46. Clinical teams respond to changes in their environment by reorganising their work processes
47. Staff have clear and shared objectives
48. There are clear processes for the involvement of service users in service development
49. There are clear criteria for identifying and establishing user involvement groups
50. Service improvement activity focuses on the patient experience of care
51. Local health and social care agencies work jointly on clinical governance issues
52. Local partnerships with health and social care agencies have clear, shared purposes
53. Staff question what they are doing, and are able to develop new and innovative models of service
54. Staff evaluate the best ways of training and learning from experience

Table 7.1.3.3.2: Validated competence level

Level 1: Low	Level 2: Medium	Level 3: High
Item No.	Item No.	Item No.
5	2	1
7	3	8
9	4	13
10	6	18
11	14	21
12	17	22
15	24	23
16	25	28
19	27	42
20	30	43
26	31	46
29	33	47
32	37	48
34	38	49
35	41	50
36	44	51
39	45	52
40		53
		54
(n) = 18	(n) = 17	(n) = 19

Conclusions

A review of clinical governance policy documents identified the major development areas and tasks required on the part of trusts to meet the clinical governance agenda. Once 'good practice' in clinical governance had been operationalised, levels of competency within the tasks were generated by an expert group, and summarised as a series of Q-sort statements. An expert group Q-sorted the statements into three levels of competency, and items were retained where there was substantial agreement between participants.

7.2 Reliability and validity of the OPGC schedule

7.2.1 Reliability (internal consistency): Factor analysis and co-efficient alpha

7.2.1.1 Factor analysis

The aim of the analysis was to assess the underlying factor structure of the OPGC, and assess the internal consistency of items within domains.

Methods

Sample size and study measures

The analysis drew on data collected during a national study of acute, ambulance and mental health / learning disabilities / social care trusts reported in section 7.3, and a detailed discussion of methodology is provided in section 7.3.2 below. Briefly, the sample consisted of 1177 completed OPGC schedules, constituting a large sample for factor analysis and regarded as ‘excellent’ by Comrey and Lee (1992). The OPGC schedules consisted of 54 items describing achievement in clinical governance, developed via literature reviews and a Q-sort study detailed above in sections 7.1.2 and 7.1.3 respectively. Items were designed as Lickert-scales in which respondents identify both the importance of each item statement as well as the extent to which it is currently being achieved on a six-point scale, from low to high.

Statistical methods

Data were analysed using SPSS-PC for Windows (SPSS Software Inc, release 9.0). A Principal Factor Analysis (PFA) model was selected in order to build a factor model based on common variance between items. While confirmatory approaches may be used to assess the validity of pre-specified factor models, they are rarely used in the initial stages of instrument development and are most suited to replication studies building on prior exploratory factor analysis (Tabachnick and Fidell 2001). The advantage of exploratory approaches in initial development is that they are able to uncover non-hypothesised cross-loadings, as well as non-hypothesised factor structures. Indeed, an exploratory factor analysis revealing predicted indicator loadings on a hypothesised factor structure constitutes a very stringent test (Kim and Mueller 1978). As items were anticipated to load on specific factors, a forced five-factor solution was specified.

Results

Factorability of the item pool

A correlation matrix of the 54 OPCG items revealed many correlations between items higher than 0.3, suggesting the likelihood of patterns in responses. Bartlett's test of Sphericity suggested that it was extremely unlikely that the sample correlations were due to chance (chi-sq. = 33714, d.f. = 1431, $p < 0.0001$). Kaiser's measure of sampling adequacy was very close to 1 (KMO MSA = 0.957), providing strong additional support for the assumption of underlying factors.

(i) *Factor extraction*

As the OPCG was designed to measure five domains, a forced five-factor solution was requested. Varimax rotation was specified in order to maximise item loadings on factors.

The scree-plot in figure 7.2.1.1 shows the first 20 eigenvalues, indicating a break at around 5 factors, accounting for 44% of the variance (table 7.2.1.1.1).

Figure 7.2.1.1 Scree-plot of factors against eigenvalues

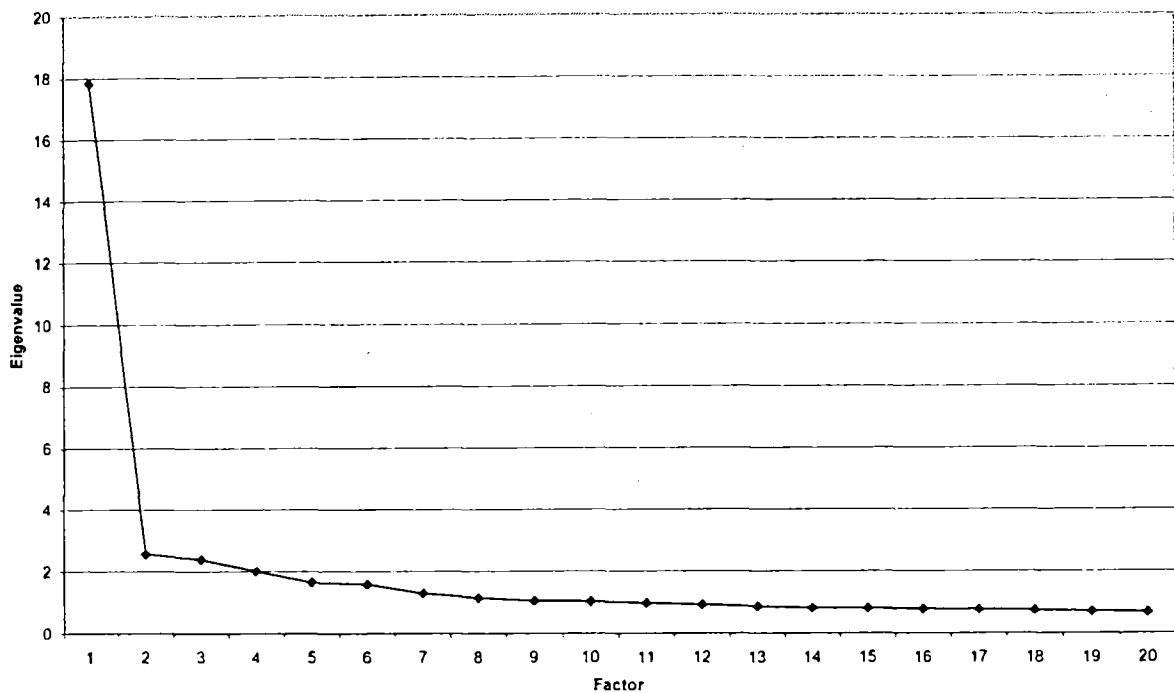


Table 7.2.1.1.1: Total variance explained, initial 5 factors, pre- and post-rotation

Factor	Initial Eigenvalues	% of Variance	Cumulative %	Rotation Sums of Squared Loadings	% of Variance	Cumulative %
	Total			Total		
1	17.827	33.012	33.012	6.516	12.066	12.066
2	2.567	4.753	37.765	6.146	11.381	23.447
3	2.401	4.446	42.211	4.425	8.194	31.641
4	2.013	3.728	45.939	4.041	7.484	39.125
5	1.649	3.054	48.992	2.723	5.042	44.167

Examination of the residual correlation matrix revealed that partial correlations between pairs of variables with the effects of factors removed are low: only 3% of residuals had values above 0.05. This suggests that there are no additional unextracted factors excluded from the five-factor model.

(ii) Model adequacy

The percentage of variance in a variable overlapping variance in factors is indicated by communalities, which are high where variables are well defined by the factor model. Table 7.2.1.1.2 shows that communalities are generally above 0.4, suggesting that the five-factor model is a good summary of the 54 variables.

(iii) Simplicity of structure and congruence with expectations

The rotated factor pattern details correlations between factors and variables. In each column (factor) there are a few high, many low and some moderate loadings, which indicated that several variables would be complex (load on more than one factor) unless a cut-off of .355 was adopted for interpretation. Under this criterion, only 4 variables load on more than one factor (table 7.2.1.1.2).

The five empirical factors match the expected OPCG domain descriptions, and items load on anticipated factors. Factor 1 contains items in domain E of the OPCG schedule associated with leadership and collaboration. Items 40 (NSF implementation) and 43 (common vision for clinical governance) are complex and also load moderately on factor 5,

corporate accountability. However, they were retained given their high factor loadings on factor 1. Factor 2 contains all of the items concerned with domain B, managing risks. Similarly, factors 3, 4 and 5 contain all of the items in domains A (quality improvement), C (improving performance) and D (corporate accountability) respectively. Consequently, the exploratory factor analysis provides good empirical support for the internal consistency of items within OPCG domains.

7.2.1.2 Co-efficient alpha

Alpha co-efficients for each of the five OPCG domains are detailed in table 7.2.1.2. Scores for each scale exceed 0.75 and thus provide additional support for the internal consistency of the OPCG domains.

Table 7.2.1.2.1: Alpha co-efficients of the OPCG domains

Domain	Alpha
Factor 1: Leadership and collaboration	0.92
Factor 2: Managing risks	0.91
Factor 3: Improving quality	0.85
Factor 4: Improving performance	0.89
Factor 5: Corporate accountability	0.78

7.2.2 Validity

7.2.2.1 Content validation

As described in section 7.1.4 above, content validation provides an empirical assessment of the extent to which scale items are a well-balanced and comprehensive sample of the domain in question. In the present study, content validation relied upon reviewing government guidance to specify the tasks (section 7.1.2), and the empirical development of

statements of achievement against those tasks using Q-sort method (section 7.1.3). This dual strategy ensured that the task areas identified in the literature were developed in the light of expert opinion.

Table 7.2.1.1.2: Factor loadings and communalities (h^2) for principal factors with varimax rotation of OPGC items

Items	Factors					h^2
	1	2	3	4	5	
52 Local partnerships with health and social care agencies have clear, shared purposes	.684	.003	.007	.003	.161	.557
51 Local health and social care agencies work jointly on clinical governance issues	.661	.006	.005	.008	.132	.535
50 Service improvement activity focuses on the patient experience of care	.588	.225	.141	.145	.007	.553
48 There are clear processes for involving service users in service development	.581	.138	.115	.185	.008	.578
49 There are clear criteria for establishing user involvement groups	.562	.172	.008	.168	.114	.618
53 Staff question what they are doing, and are able to develop new and innovative models of service	.540	.262	.282	.222	-.003	.578
47 Staff have clear and shared objectives	.512	.279	.236	.319	.008	.557
40 NSF implementation is integrated with business planning and quality improvement programmes	.508	.220	.191	.009	.376	.666
46 Clinical teams respond to changes in their environment by reorganising their work processes	.497	.262	.306	.200	-.005	.535
54 Staff evaluate the best ways of training and learning from experience	.489	.234	.324	.305	-.002	.618
33 Service delivery plans include quality improvement activity	.485	.263	.183	.201	.245	.458
38 Decisions about service developments or cutbacks are made on clear criteria	.469	.233	.197	.004	.250	.445
43 The organisation shares a common vision for clinical governance	.463	.308	.246	.120	.391	.573
45 Clinical teams receive performance feedback	.453	.308	.299	.179	.203	.553
39 Local and national priorities from NSFs and HimPs are used to prioritise service development	.427	.157	.157	.003	.322	.550
42 Clinical protocols are shared with staff who work outside this organisation	.398	.205	.271	.167	.134	.326
41 External guidelines are critically appraised before local adoption or development	.396	.204	.319	.136	.145	.362
44 Leadership skills are identified and developed through leadership training programmes	.383	.234	.008	.332	.220	.392
1 Teams work across boundaries to address quality issues	.367	.293	.323	.114	.008	.358
17 Clear action plans are developed in response to adverse incidents	.197	.734	.112	.119	.149	.653
16 Adverse incident data is regularly reviewed and discussed	.120	.708	.116	.120	.192	.592
18 Following identification of a problem from adverse events data, clinical quality is improved	.253	.677	.277	.194	.004	.645
13 Following identification of a problem from risk data, clinical quality is improved	.232	.666	.287	.122	.003	.625
14 Clear action plans are developed in response to identified clinical risks	.259	.653	.176	.136	.160	.605
12 Risk management data is regularly reviewed and discussed	.159	.570	.188	.009	.223	.484
15 Staff are trained to use the risk management system	.181	.522	.166	.134	.187	.560
21 Following identification of a problem from complaints, clinical quality is improved	.293	.506	.256	.149	.003	.506
19 Staff are trained to use the adverse events systems	.009	.505	.220	.158	.160	.509
20 Complaints are collated	.141	.452	.008	.008	.189	.319
7 Clinical issues are raised for discussion	.143	.407	.355	.113	.215	.481
22 There is a 'no-blame' culture around reporting adverse events and near misses	.245	.355	.159	.152	.168	.454
23 Staff can raise clinical concerns about their colleagues in confidence	.231	.316	.227	.175	.201	.514
8 Discussions on clinical issues are not dominated by any single profession	.188	.283	.698	.004	.154	.320

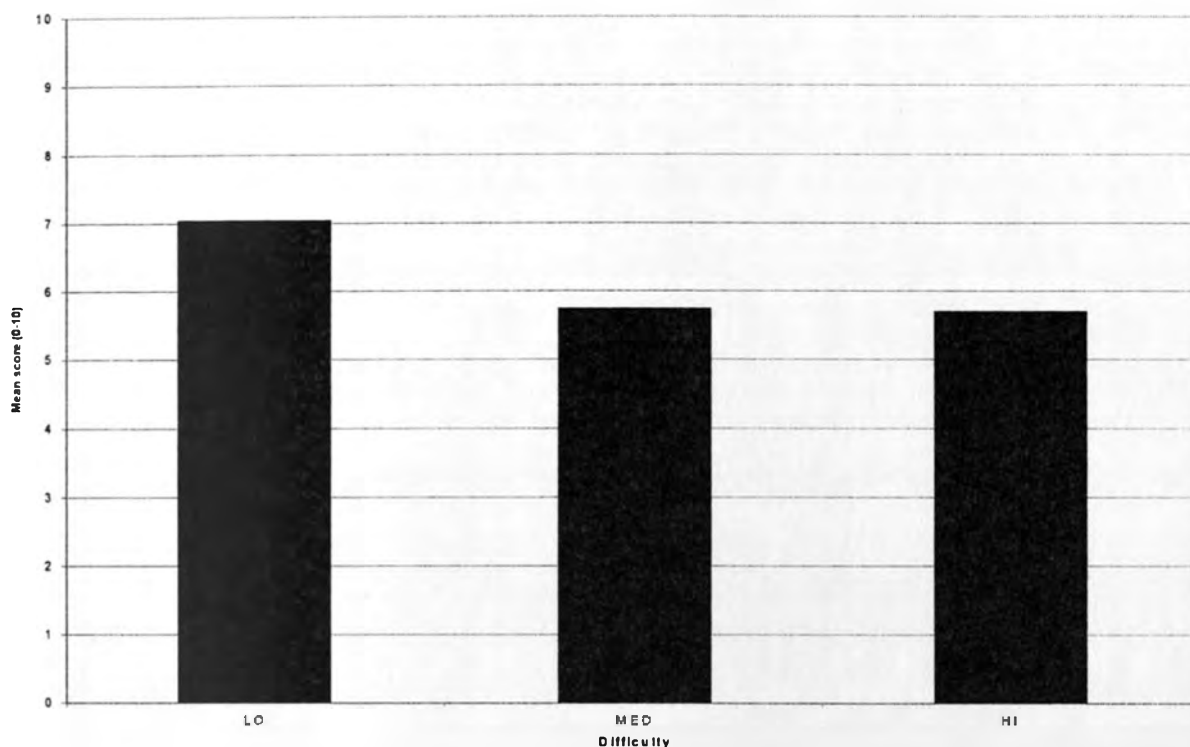
Items	Factors					
	1	2	3	4	5	h ²
6 Research information is used consistently to inform our approach to quality improvement	.172	.152	.692	.123	.008	.568
4 Research evidence is regularly reviewed and discussed	.105	.139	.628	.122	.009	.480
5 Training in evidence-based practice and critical appraisal is available to staff	.134	.118	.539	.213	.114	.476
3 Where appropriate, staff modify their care processes to reflect the best practice of others	.226	.266	.490	.134	-.004	.429
25 Clinical indicators are used to reflect on, review and integrate services	.408	.290	.451	.006	.176	.716
24 There is good access to agreed clinical indicators	.308	.268	.418	.007	.165	.644
11 Topics for audit are selected according to their potential impact on care quality	.230	.254	.410	.002	.169	.317
10 Training in clinical audit is available to staff	.007	.150	.404	.238	.250	.405
2 Staff identify best practice through benchmarking against other providers	.248	.228	.390	.009	.002	.319
9 Staff participate in clinical audit activity	.003	.296	.373	.230	.249	.410
31 Staff appraisal is used as an opportunity to reflect on progress and plan for future development	.160	.156	.124	.774	.152	.706
29 There is an annual appraisal process for most staff	.008	.157	.009	.770	.158	.691
26 Staff have development plans which identify training and development opportunities	.177	.129	.198	.718	.175	.635
30 An agreed work and development programme is used as the basis of staff appraisal	.259	.135	.173	.700	.131	.665
27 Training identified in staff development plans matches individual needs to organisational needs	.288	.215	.224	.645	.134	.642
28 New skills gained through development activity are used in clinical settings	.329	.233	.322	.423	.011	.478
35 There is a formal clinical governance committee, reporting to the board	.007	.175	.007	.141	.586	.446
36 There are local arrangements to collate information for the clinical governance committee	.190	.287	.146	.162	.541	.540
32 There is an executive director with responsibility for developing the clinical governance agenda	.202	.134	.002	.141	.502	.348
34 Clinical areas have a nominated clinical governance lead	.227	.230	.215	.177	.434	.394
37 Organisation-wide clinical governance systems are underpinned by systems in clinical areas	.372	.357	.254	.202	.425	.710

7.2.2.2 Construct validation

As there is no 'gold standard' pre-existing measure of clinical governance progress, it was not possible to assess the criterion validity of the OPCG schedule by comparing scores between the two measures. Consequently, recourse was made to construct validation i.e. the ability of the OPCG to confirm expected hypotheses. As the Q-sort expert group were asked to rate OPCG items as indicative of 'low', 'medium' or 'high' achievement, one would expect that items classified as 'low' would be scored more highly than those classified as 'medium' would. Similarly, one would expect those classified as 'medium' to be scored more highly than those classified as 'high'. As the number of items within each category varied between 17 and 19, an aggregate score for each of the three categories was obtained summing the items and standardising the final score between 0-10 using the formula: $\text{Score} = (\text{sum of items} - (\text{no. of items}) / 5 * 10)$

Figure 7.2.2.2.1 below shows mean scores and 95% confidence intervals for items within each of the three levels of difficulty. Items reflecting low achievement received a higher mean score than those items indicative of as 'medium' ($t = 34.31, p < .000$) or 'high' difficulty ($t = 35.378, p < .000$), as hypothesised. However, while mean score for 'medium' items was higher than for 'high' items, the difference was not statistically significant ($t = .931, p = .352$). This lends some empirical support to the assignation of OPCG items to levels of attainment, as the respondents indicated highest achievement against items regarded as most easy to achieve. The inability to empirically separate scores for medium and high difficult items may be due to either inappropriate assignation, or to the failure of trusts to move beyond the 'easy' agenda.

Figure 7.2.2.2.1: Construct validation of levels of difficulty: Mean achievement scores for items indicative of low, medium and high achievement



	Mean	Std. Error	Std. Deviation	Lower 95% CI	Higher 95% CI
LO	7.04	0.04	1.30	6.97	6.11
MED	5.74	0.04	1.44	5.66	5.82
HI	5.70	0.04	1.44	5.62	5.78

(n = 1177).

Conclusions

The internal consistency of OPCG items within five anticipated domains of improving quality, managing risks, improving performance, corporate accountability and leadership and collaboration was examined using a forced five-factor exploratory principal factor

analysis. After rotation, the model explained 44% of the common variance, and residual analysis suggested no further unextracted factors. The five empirical factors matched the expected domains, and items loaded on their expected domains with only 4 items loading on more than 1 factor above a .355 threshold. All five domains had co-efficient alpha values between 0.78-0.92, which additionally suggests that items aggregated into the five domains are internally consistent. The empirical development work outlined in section 7.1 lends support for the content validity of the schedule, and its ability to discriminate between achievement on aggregations of items previously designated as indicative of 'low' difficulty from those designated 'medium' and 'high' difficulty lends qualified support to its construct validity.

CHAPTER 8

APPLYING THE OPCG COMPETENCY FRAMEWORK: A NATIONAL STUDY OF PROGRESS IN CLINICAL GOVERNANCE

Introduction

The early empirical studies of clinical governance implementation reviewed in chapter 2 assessed progress largely in structural terms. This chapter applies the OPCG framework developed above to compare managers' perceptions of progress in clinical governance against a range of clinical governance competencies. These competencies related to the existence of structures, the operation of processes and their effects (outcomes) on the way that quality assurance and improvement work is undertaken. The study was undertaken five years after the UK government outlined clinical governance as part of its 1997 health care quality reforms, at the mid-point of a ten-year programme. It comprised a stratified sample of acute, ambulance and mental health / learning disabilities trusts in England (n=100), from each of which up to ten board- and ten directorate-level managers were randomly sampled. The Organisational Progress in Clinical Governance (OPCG) schedule, outlined in section 6.1 and 6.2 above, was used to explore respondent's perceptions of the importance of, and their own organisation's achievement in, 54 clinical governance competencies. Competency scores were further aggregated into five domains: improving quality; managing risks; improving staff performance; corporate accountability; and leadership and collaboration. For both competency items and aggregated domains, the difference between ratings of importance and achievement was termed a shortfall.

Of 1916 managers surveyed across 100 trusts, 1177 (61.4%) responded. The competencies considered most important and recording highest levels of achievement related to corporate accountability structures and clinical risks. The highest shortfalls between importance and achievement were reported in joint working across local health communities, feedback of performance data, and user involvement. When aggregated into domains, greatest achievement was shown in the assurance related areas of corporate accountability and risk management, with considerably less achievement and consequently higher shortfalls in quality improvement and leadership & collaboration. When managers' perceptions of progress in clinical governance at trusts before, during and after a CHI clinical governance review were compared, few differences were found suggesting that the impact of CHI visits may be slight, temporary, and limited to performance improvement and corporate accountability domains. It appears that five years into the programme, structures and systems for clinical governance are have been established. Yet, there seems to be more progress in areas concerned with quality assurance than quality improvement. Results also suggested greater progress in competencies related to structures than those related to processes or outcomes .

8.1 Methodology

The study employed a structured postal questionnaire to explore perceptions of progress in clinical governance among board and directorate managers in English acute, ambulance and mental health / learning disabilities trusts. Primary care organisations were specifically

excluded, given large-scale reorganisation within this sector occasioned by the move from Primary Care Groups to Primary Care Trusts during April 2002.

Sampling

Identification of all eligible trusts was a potentially daunting prospect, given the scale and pace of organisational restructuring within the UK NHS. Fortunately, the National Audit Office (NAO) established a sample frame of all eligible English acute, ambulance and mental health / learning disabilities (MH/LD) trusts as part of a larger Value For Money (VFM) study of clinical governance in the NHS. In order to ensure representation of different sectors within the sample, a stratified random sample of 100 trusts was drawn from the sample frame, reflecting the proportion of each type of organisation. Consequently, 68 acute, 11 ambulance and 21 MH/LD trusts were randomly selected for inclusion. To gather perceptions from multiple respondents, up to 10 board members and 10 directorate managers at each trust in the study were randomly sampled from Binleys' database of trust managers (Binleys 2002). Where fewer than 10 such staff members were listed, all eligible staff were included. In total, 1916 participants across 100 trusts were selected. Participants received a six-page questionnaire, re-mailed to non-respondents at four and eight weeks (appendix 12).

Study measures

Developed at the University of Birmingham Health Services Management Centre and detailed in sections 6.1 and 6.2 above, the Organisational Progress in Clinical Governance (OPCG) schedule assesses respondent perceptions of achievement in a series of

organisational competencies related to clinical governance. Briefly, the OPCG consists of 54 statements under five domains - improving quality; managing risk; improving staff performance; corporate accountability; and leadership and collaboration - allowing the aggregation of competency item scores into domain scores. As these domains consist of different numbers of items, reported domain scores are standardised between 0-10 to aid comparison.

Analysis addressed the following questions:

1. Across all domains, which competencies are considered most important, show most achievement, or have shortfalls (in which the ratings for importance exceed the ratings for achievement)?
2. Within each domain, which clinical governance competencies are considered most important, show achievement, or have shortfalls?
3. When competency items are aggregated into domains, which domains show most achievement, and where are the shortfalls?
4. What progress is there in clinical governance structures, processes and outcomes?
5. Do perceived levels of achievement vary between individual trusts?
6. What impact do CHI visits, trust type and respondent's position in the organisation have on perceptions of achievement in OPCG domains, structures/ processes and outcomes, and aggregates of competencies at high / medium / and low levels?

Statistical methods

Data were analysed using SPSS-PC for Windows software (SPSS Software Inc, release 9.0). To explore the effects of trust type, respondent's status and CHI visit-status on OPCG domain scores, structure/ process/ outcome aggregates, and low / medium / high competency aggregates, three 3 x 3 x 2 between-subjects multivariate analysis of variance (MANOVA) analyses were performed. The first used OPCG domain scores as dependent variables, the second structure/ process/ outcome aggregates as dependent variables, and the third low / medium / high competency aggregates as dependent variables. For each model, correlations between multiple dependent variables were assessed using Pearson product-moment correlation. Wilkes' lambda and Pillai's trace was used to assess the main effect of independent variables on combined dependent variables. For significant main effects, relationships between independent and dependent variables were explored further using univariate F-tests and post-hoc Scheffe tests.

Data cleaning and preparation for multivariate analysis

The quality of the analysis was improved by adopting data cleaning procedures and assessing missing data. While normal distribution of dependent variables was not an issue given the large sample size, data was further screened for multivariate outliers, and the assumption of homogeneity of variances was tested prior to MANOVA analysis (Tabachnick and Fidell 2001).

(a) Data cleaning

Data accuracy for all 1177 cases was assessed through the SPSS FREQUENCIES command for categorical data, revealing three data points to be out of range. These were checked against original paper records and amended. Subsequent entry validation, assessed by an examination of the first 118 (10%) cases of entered data against the original paper questionnaires, showed that all data points had been correctly entered.

(b) Assessment of missing data

SPSS FREQUENCIES revealed no variables with missing data, which suggests that respondents found the schedule very easy to complete.

8.2 Results

Of 1916 eligible study participants, 1177 (61.4%) returned completed questionnaires.

1. Across all domains, which competencies are perceived as most important / show highest achievement / possess the highest shortfalls?

(a) Importance

Overall, respondents rated items highly for importance – the mean item rating was 5.1, and no single item was rated lower than 4.5. Figure 8.2.1 shows means for competency items with the 15 lowest / highest importance scores, and means, standard deviations and 95% confidence intervals appear in table 8.2.6. Items regarded as most important included formal committee structure (35, 32); 'no blame' culture (22); raising concerns and action planning / improving quality around risks and adverse events (13-14, 16-18, & 23). Least

important items included benchmarking (2); use of research and clinical indicators (4-6, 25); and joint working with partner agencies and shared protocols (51-52, 41).

Figure 8.2.1: Perceived importance of competencies

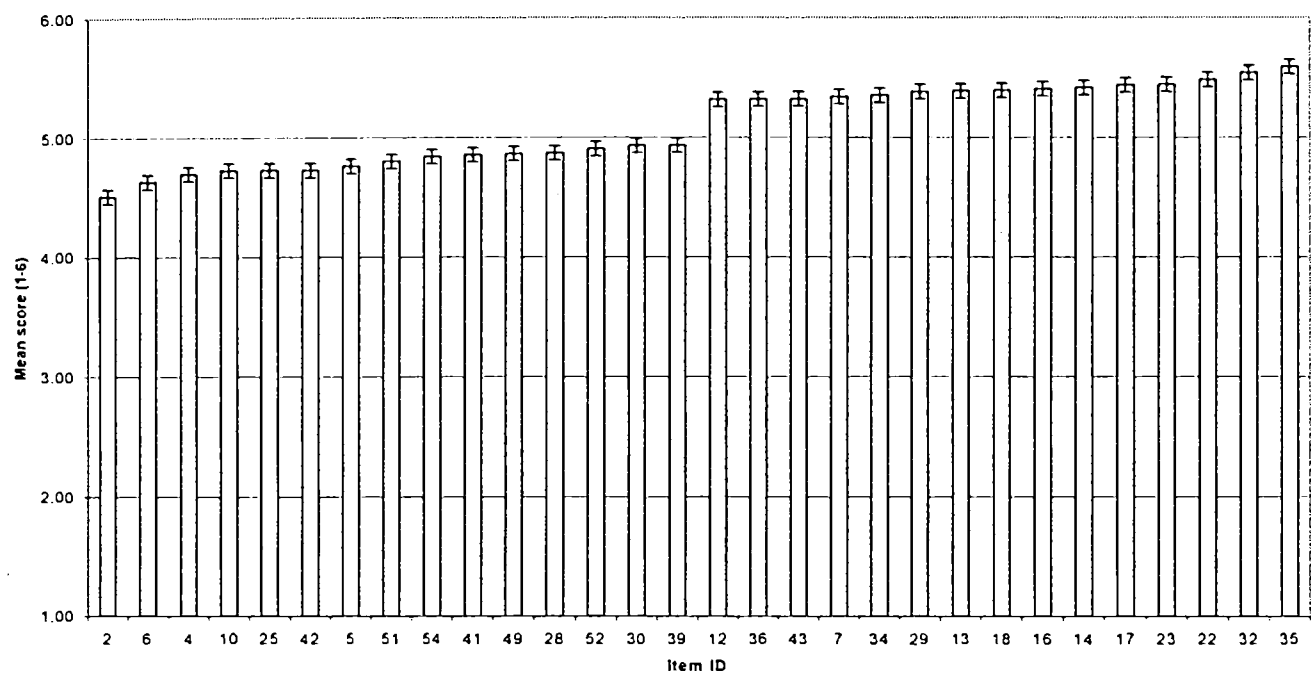


Table 8.2.1: Perceived importance of competencies

Item no.	Mean	95%CI (+/-)	Std Deviation
2	4.51	0.06	0.93
6	4.63	0.06	0.95
4	4.7	0.06	0.94
10	4.72	0.06	0.95
42	4.72	0.06	1.01
25	4.73	0.06	0.94
5	4.76	0.06	0.94
51	4.8	0.06	1.02
54	4.84	0.06	0.91
41	4.85	0.06	0.9
28	4.87	0.06	0.87
49	4.87	0.06	1.01
52	4.91	0.06	0.99
30	4.93	0.06	0.95
38	4.94	0.06	0.97
12	5.32	0.04	0.77
36	5.32	0.04	0.83
43	5.32	0.04	0.83
7	5.34	0.04	0.74
34	5.35	0.04	0.84
13	5.39	0.04	0.75
18	5.39	0.04	0.73
29	5.39	0.04	0.81
16	5.41	0.04	0.75
14	5.42	0.04	0.7
17	5.44	0.04	0.73
23	5.44	0.04	0.74
22	5.48	0.04	0.75
32	5.54	0.04	0.79
35	5.59	0.04	0.74

(b) Achievement

In order to facilitate comparisons between competencies in different domains, figure 8.2.2 shows means for competency items with the 15 lowest / highest achievement scores.

Respondents perceived higher achievement against items concerning structural change for corporate accountability than against those concerned with quality improvement or

collaborative outcomes. High scoring items included committee structures (35, 32); collation of complaints / information (20, 36); raising clinical issues (7); discussing risk and adverse event data (16) and priority planning (39, 40). Conversely, low scoring items included joint work across local health communities (51-52); benchmarking for quality improvement (2); use of research evidence (6, 4); using clinical indicators (24-25); user involvement (49) and promoting clinical teams (45-46).

Figure 8.2.2: Achievement in competencies

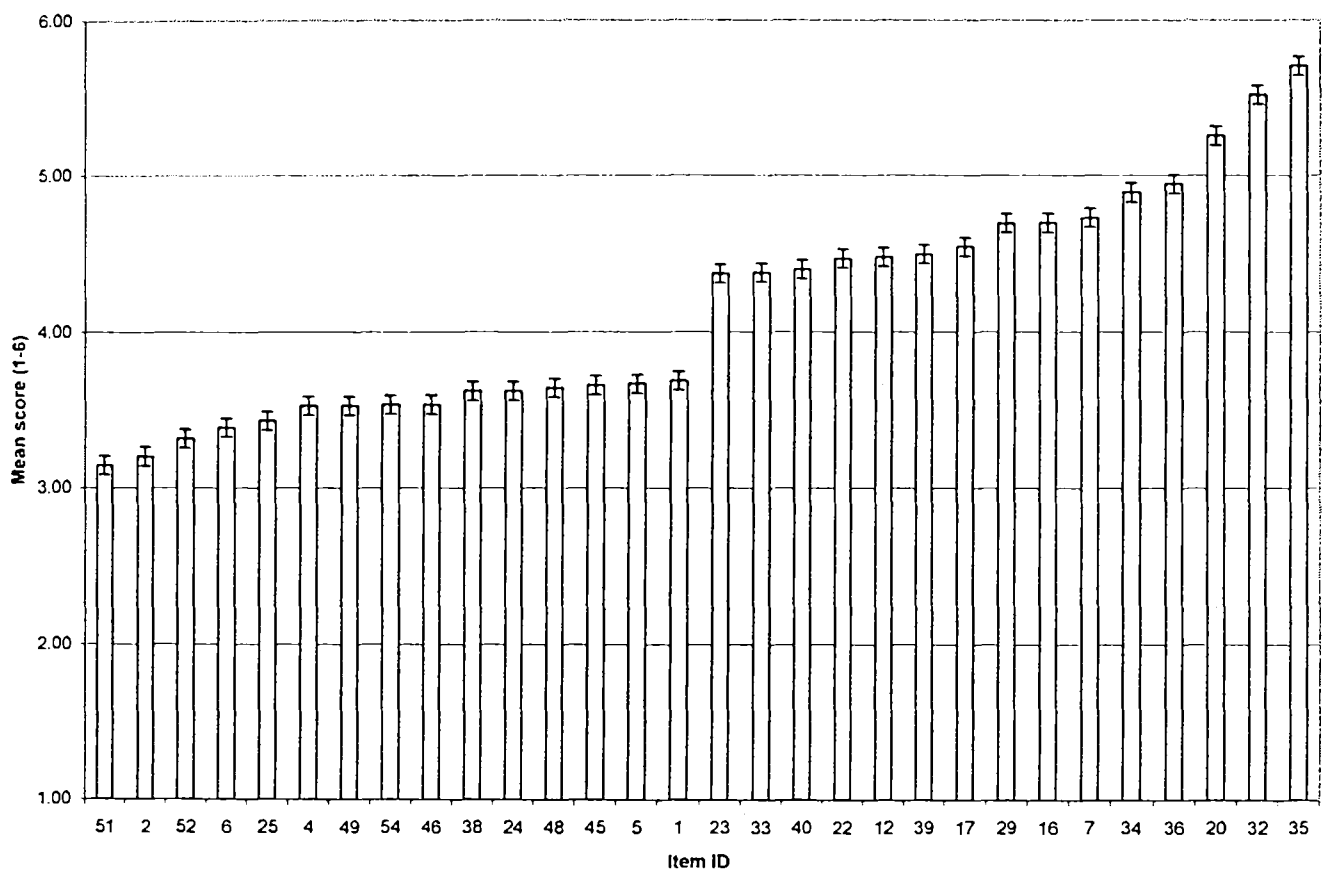


Table 8.2.2: Achievement in competencies

Item no.	Mean	95%CI (+/-)	Std Deviation
51	3.14	0.08	1.29
2	3.2	0.06	1.04
52	3.32	0.08	1.29
6	3.38	0.06	1.11
25	3.42	0.06	1.1
4	3.52	0.08	1.21
49	3.52	0.08	1.29
54	3.52	0.06	1.04
46	3.53	0.06	1.08
24	3.61	0.06	1.17
38	3.62	0.08	1.22
48	3.64	0.08	1.26
45	3.65	0.08	1.21
5	3.66	0.08	1.32
1	3.69	0.06	0.99
23	4.37	0.06	1.15
33	4.38	0.06	1.13
40	4.41	0.06	1.12
22	4.47	0.06	1.2
12	4.48	0.06	1.19
39	4.51	0.06	1.09
17	4.55	0.06	1.12
16	4.7	0.06	1.14
29	4.7	0.08	1.25
7	4.73	0.06	0.97
34	4.9	0.08	1.32
36	4.95	0.06	1.13
20	5.25	0.06	0.93
32	5.52	0.06	0.94
35	5.7	0.04	0.7

(c) Shortfalls

Shortfalls (figure 8.2.3) were highest in joint work across local health communities (51-52); clarity of service development criteria (38); performance feedback and clear objectives (45, 47); working across boundaries and reorganising work processes (1, 46) and user involvement (48-50).

Figure 8.2.3: Shortfalls in competencies

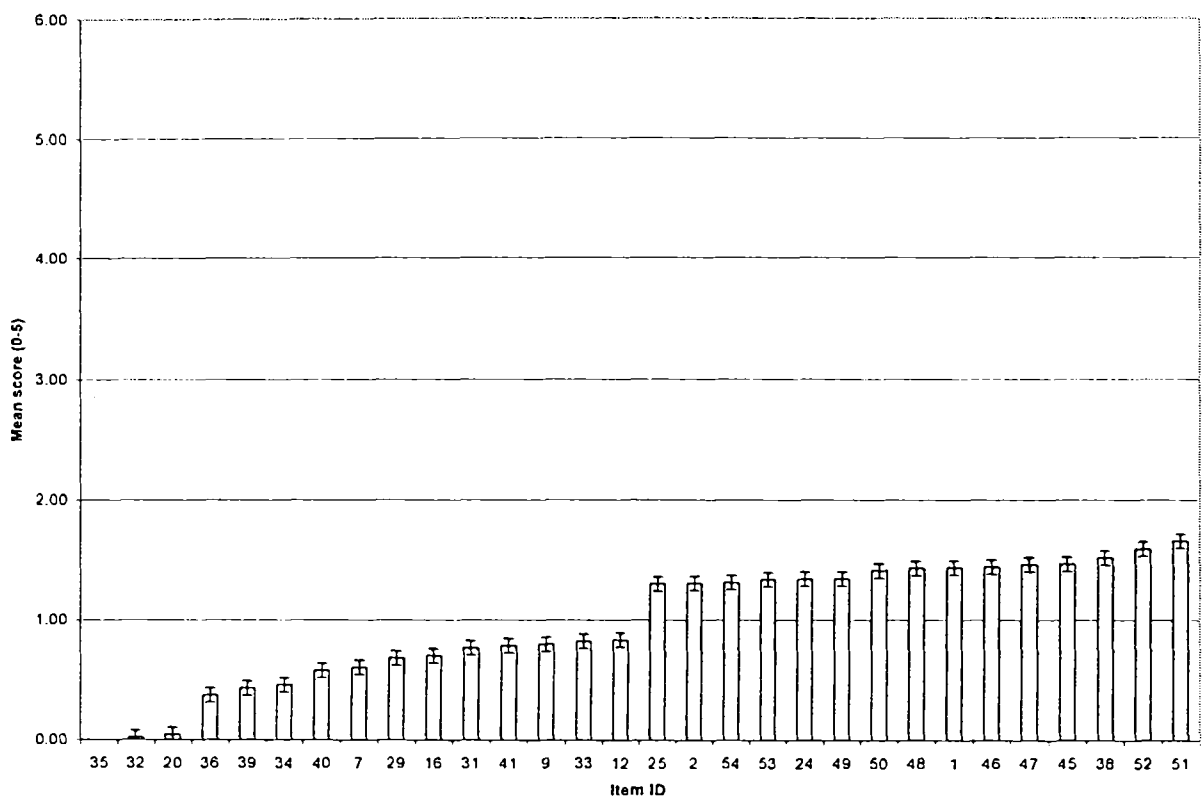


Figure 8.2.3 shows the 15 items with the lowest and 15 with the highest, shortfalls between importance and achievement scores. Those with the highest shortfalls required most urgent attention.

Table 8.2.3: Shortfall in competencies

Item no.	Mean	95%CI (+/-)	Std Deviation
35	-0.11	0.04	0.75
32	0.02	0.06	0.95
20	0.05	0.06	0.89
36	0.37	0.06	1.02
39	0.43	0.06	1.20
34	0.46	0.08	1.21
40	0.58	0.06	1.12
7	0.61	0.06	0.91
29	0.69	0.08	1.25
16	0.71	0.06	1.07
31	0.77	0.06	1.17
41	0.78	0.06	1.10
9	0.80	0.06	1.09
33	0.82	0.06	1.06
12	0.84	0.06	1.15
25	1.31	0.06	1.16
2	1.31	0.06	1.17
54	1.32	0.06	1.10
53	1.34	0.06	1.11
24	1.35	0.08	1.21
49	1.35	0.08	1.30
50	1.41	0.06	1.19
48	1.43	0.08	1.26
1	1.44	0.06	1.11
46	1.44	0.06	1.17
45	1.47	0.08	1.23
47	1.47	0.06	1.16
38	1.51	0.08	1.32
52	1.59	0.08	1.30
51	1.66	0.08	1.29

2. Within each domain, which clinical governance competencies are considered most important, show achievement, or have shortfalls?

Quality improvement

Eleven OPCG items concern aspects of quality improvement activity including benchmarking practice, using research evidence, auditing practice and clinical indicators.

Importance

Benchmarking, using research evidence and training in critical appraisal & audit were seen as slightly less important than participating in audit activity.

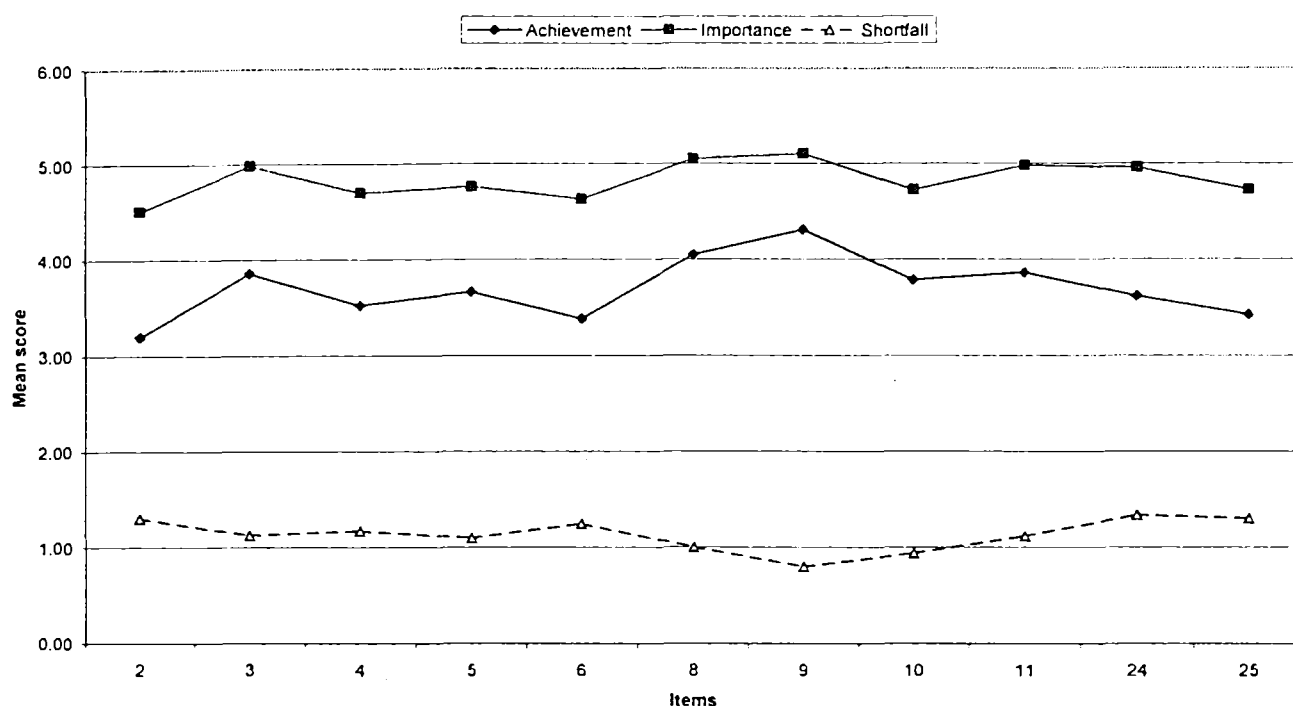
Achievement

There was only moderate achievement against selection of audit topics based on quality improvement considerations.

Shortfall

Principal shortfalls included using research to inform quality improvement and selection of audit topics based on quality improvement needs.

Figure 8.2.4: Quality improvement items: importance / achievement / shortfall



Quality improvement items	Means		
	Achievement	Importance	Shortfall
2 Staff identify best practice through benchmarking against other providers	3.20	4.51	1.31
3 Where appropriate, staff modify their care processes to reflect the best practice of others	3.86	4.99	1.13
4 Research evidence is regularly reviewed and discussed	3.52	4.70	1.17
5 Training in evidence-based practice and critical appraisal is available to staff	3.67	4.76	1.10
6 Research information is used consistently to inform our approach to quality improvement	3.38	4.63	1.25
8 Discussions on clinical issues are not dominated by any single profession	4.05	5.05	1.00
9 Staff participate in clinical audit activity	4.31	5.11	0.80
10 Training in clinical audit is available to staff	3.79	4.73	0.94
11 Topics for audit are selected according to their potential impact on care quality	3.86	4.98	1.12
24 There is good access to agreed clinical performance indicators	3.62	4.96	1.34
25 Clinical indicators are used to reflect on, review and integrate services	3.43	4.73	1.30

Table 8.2.4: Quality improvement achievement data

Frequency tables for quality improvement items						
	1.00	2.00	3.00	4.00	5.00	6.00
	Count	Count	Count	Count	Count	Count
A2	36	276	423	316	109	17
A3	7	68	319	505	247	31
A4	37	230	310	330	219	51
A5	60	190	261	328	249	89
A6	45	215	375	352	164	26
A8	31	104	219	368	335	120
A9	11	67	190	356	387	166
A10	59	151	266	327	254	120
A11	23	120	286	386	288	74
A24	43	160	318	396	201	59
A25	47	189	364	395	155	27

Frequency tables for quality improvement items (%ages)						
	1.00	2.00	3.00	4.00	5.00	6.00
	%	%	%	%	%	%
A2	3.1%	23.4%	35.9%	26.8%	9.3%	1.4%
A3	.6%	5.8%	27.1%	42.9%	21.0%	2.6%
A4	3.1%	19.5%	26.3%	28.0%	18.6%	4.3%
A5	5.1%	16.1%	22.2%	27.9%	21.2%	7.6%
A6	3.8%	18.3%	31.9%	29.9%	13.9%	2.2%
A8	2.6%	8.8%	18.6%	31.3%	28.5%	10.2%
A9	.9%	5.7%	16.1%	30.2%	32.9%	14.1%
A10	5.0%	12.8%	22.6%	27.8%	21.6%	10.2%
A11	2.0%	10.2%	24.3%	32.8%	24.5%	6.3%
A24	3.7%	13.6%	27.0%	33.6%	17.1%	5.0%
A25	4.0%	16.1%	30.9%	33.6%	13.2%	2.3%

Risk management

Thirteen OPCG items concern aspects of risk management, including the review and discussion of data on risk and clinical incidents / near misses; planning and service development in the light of identified trends; and a blame-free reporting culture.

Importance

There was little variation in perceived importance; all items gained a mean score of more than five indicating high importance.

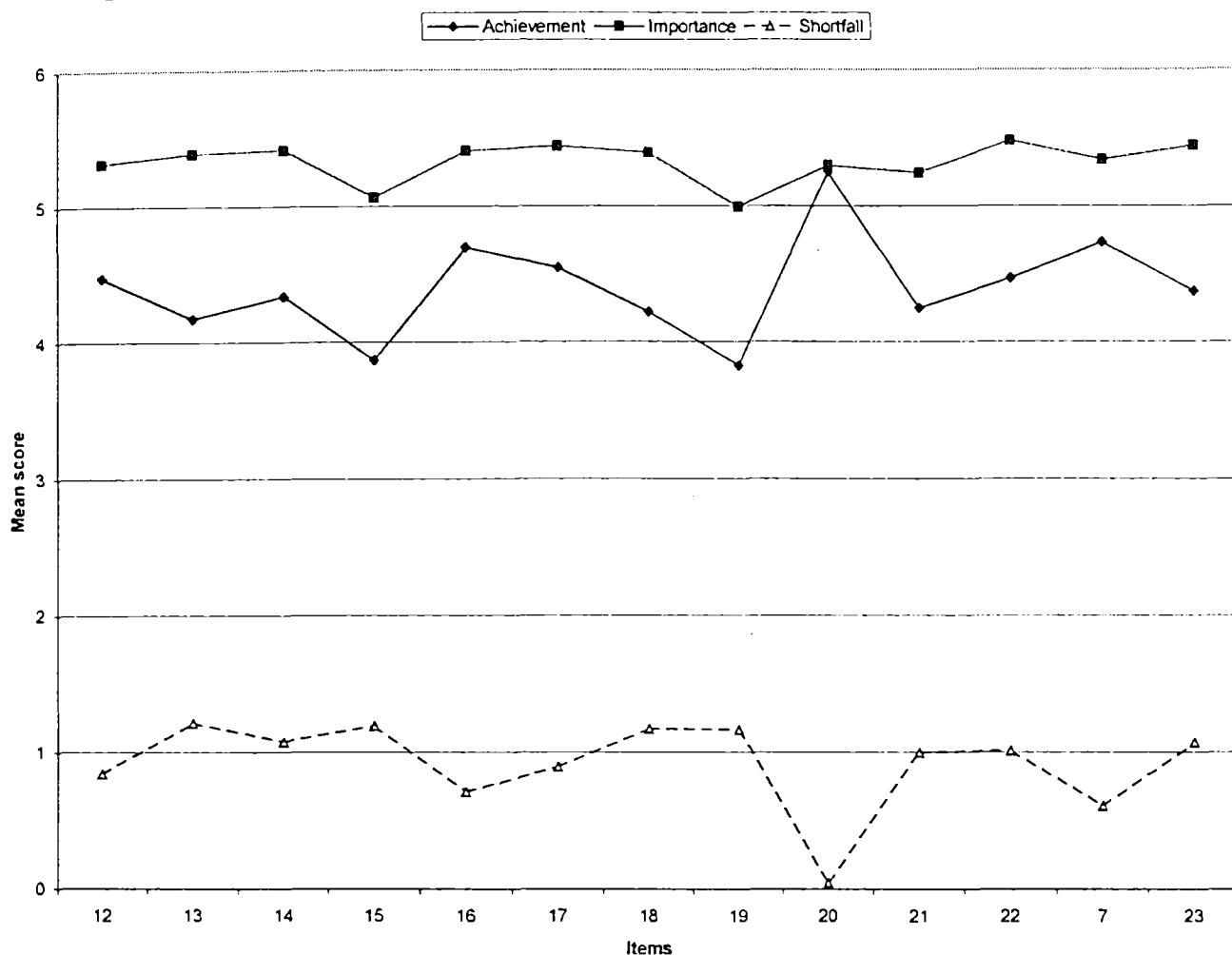
Achievement

Moderate scores (3-4) were achieved for training in use of risk management and adverse event systems, pointing to the need to ensure that staff are aware of reporting systems and are able to use them appropriately. While organisations scored highly on collation of complaints and reviewing and discussing data on adverse events and risk management, they scored less well on improving quality once problems are identified from data (items 13,18 and 21).

Shortfall

Areas with greatest perceived shortfall between importance and achievement concern training in adverse event and risk management systems, and moving from identification of problems in adverse event / risk data to action to improve quality.

Figure 8.2.5: Risk management items: importance / achievement / shortfall



Risk management items	Means		
	Achievement	Importance	Shortfall
12 Risk management data is regularly reviewed and discussed	4.48	5.32	0.84
13 If problems are identified from risk data, quality is improved	4.18	5.39	1.21
14 Clear action plans are developed in response to identified clinical risks	4.34	5.42	1.08
15 Staff are trained to use the risk management system	3.87	5.07	1.20
16 Adverse incident data is regularly reviewed and discussed	4.70	5.41	0.70
17 Clear action plans are developed in response to adverse incidents	4.55	5.44	0.89
18 If problems are identified from adverse events data, quality is improved	4.22	5.39	1.17
19 Staff are trained to use adverse event systems	3.84	5.00	1.16
20 Complaints are collated	5.25	5.30	0.05
21 If problems are identified from complaints data, quality is improved	4.25	5.25	1.00
22 There is a 'no-blame' culture for reporting adverse events / near misses	4.47	5.48	1.01
7 Clinical issues are raised for discussion	4.73	5.34	0.61
23 Staff can raise clinical concerns about their colleagues in confidence	4.38	5.44	1.07

Table 8.2.5: Risk management achievement data

Frequency tables for risk management items						
	1.00	2.00	3.00	4.00	5.00	6.00
	Count	Count	Count	Count	Count	Count
A12	18	62	152	285	422	238
A13	12	66	206	415	375	103
A14	9	55	188	376	375	174
A15	52	122	257	355	278	113
A16	11	42	117	263	427	317
A17	12	52	118	332	421	242
A18	7	57	197	423	393	100
A19	34	135	276	377	254	101
A20	5	14	38	140	402	578
A21	9	44	208	423	366	127
A22	24	59	136	317	399	242
A7	3	24	99	276	529	246
A23	20	48	166	382	356	205

Frequency tables for risk management items (%ages)						
	1.00	2.00	3.00	4.00	5.00	6.00
	%	%	%	%	%	%
A12	1.5%	5.3%	12.9%	24.2%	35.9%	20.2%
A13	1.0%	5.6%	17.5%	35.3%	31.9%	8.8%
A14	.8%	4.7%	16.0%	31.9%	31.9%	14.8%
A15	4.4%	10.4%	21.8%	30.2%	23.6%	9.6%
A16	.9%	3.6%	9.9%	22.3%	36.3%	26.9%
A17	1.0%	4.4%	10.0%	28.2%	35.8%	20.6%
A18	.6%	4.8%	16.7%	35.9%	33.4%	8.5%
A19	2.9%	11.5%	23.4%	32.0%	21.6%	8.6%
A20	.4%	1.2%	3.2%	11.9%	34.2%	49.1%
A21	.8%	3.7%	17.7%	35.9%	31.1%	10.8%
A22	2.0%	5.0%	11.6%	26.9%	33.9%	20.6%
A7	.3%	2.0%	8.4%	23.4%	44.9%	20.9%
A23	1.7%	4.1%	14.1%	32.5%	30.2%	17.4%

Performance improvement

Six items concern performance improvement, including professional development and appraisal.

Importance

Respondents considered most items important with a mean of approximately five. Items concerning annual appraisal were slightly more important than the use of new skills gained through staff development.

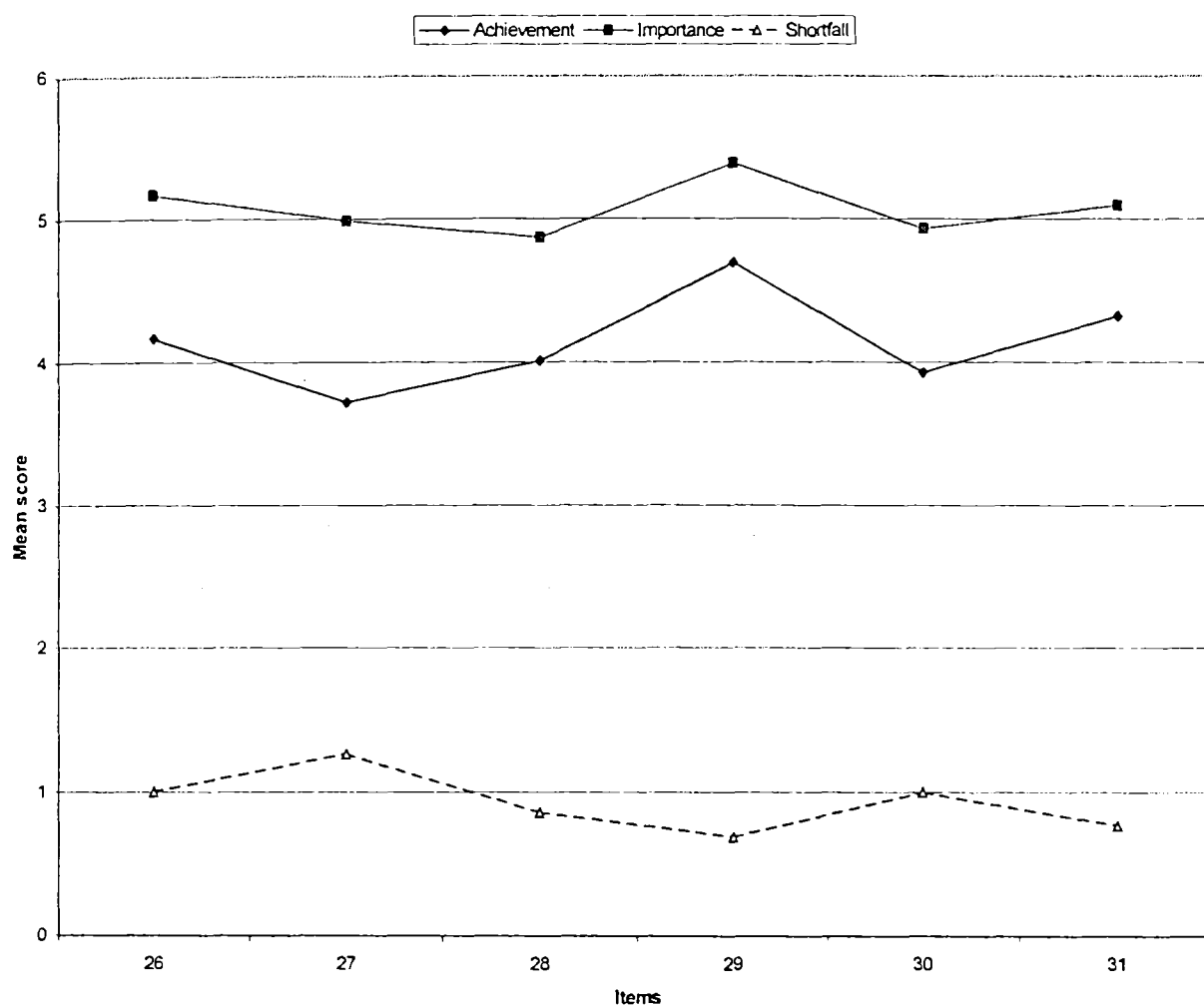
Achievement

Only moderate achievement scores (3-4) were attained for the matching of individual and organisational needs through staff development. The existence of staff appraisal systems and their use to identify development was scored higher, between 4-5.

Shortfall

The largest perceived shortfalls concerned staff development plans identifying training opportunities; and matching individual and organisational needs in staff development.

Figure 8.2.6: Performance improvement items: importance / achievement / shortfall



Performance improvement items	Means		
	Achievement	Importance	Shortfall
26 Staff development plans identify training and development opportunities	4.18	5.17	1.00
27 Staff development training matches individual and organisational needs	3.73	4.99	1.26
28 New skills gained through development activity are used in clinical settings	4.02	4.87	0.85
29 There is an annual staff appraisal process for most staff	4.70	5.38	0.68
30 Agreed work and development programmes form the basis of staff appraisal	3.93	4.93	1.00
31 Staff appraisal is used to reflect on progress and plan future development	4.33	5.10	0.77

Table 8.2.6: Performance improvement achievement data

Frequency tables for performance improvement items						
	1.00	2.00	3.00	4.00	5.00	6.00
	Count	Count	Count	Count	Count	Count
A26	19	74	212	393	334	145
A27	32	136	323	391	219	76
A28	13	66	248	480	299	71
A29	33	42	113	224	410	355
A30	43	112	229	400	281	112
A31	33	52	179	330	401	182

Frequency tables for performance improvement items (%ages)						
	1.00	2.00	3.00	4.00	5.00	6.00
	%	%	%	%	%	%
A26	1.6%	6.3%	18.0%	33.4%	28.4%	12.3%
A27	2.7%	11.6%	27.4%	33.2%	18.6%	6.5%
A28	1.1%	5.6%	21.1%	40.8%	25.4%	6.0%
A29	2.8%	3.6%	9.6%	19.0%	34.8%	30.2%
A30	3.7%	9.5%	19.5%	34.0%	23.9%	9.5%
A31	2.8%	4.4%	15.2%	28.0%	34.1%	15.5%

Corporate accountability

Five items concern corporate accountability, including systems and structures for clinical governance.

Importance

Items concerning systems and structures were considered to be highly important, with mean scores above five.

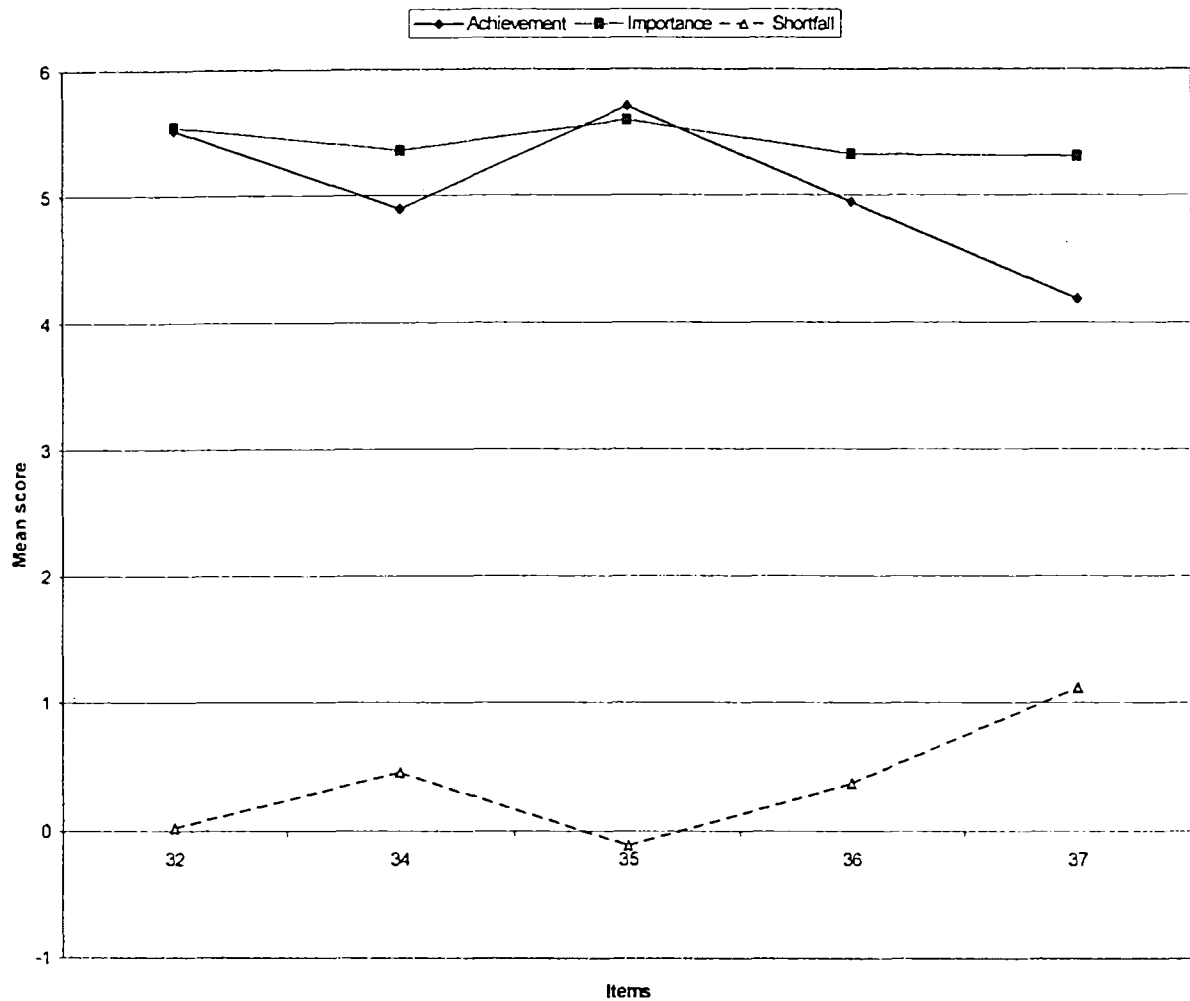
Achievement

Understandably, the early targets of the existence of an executive committee and an executive director with lead responsibility scored more than five. However, respondents rated achievement in the underpinning of organisation-wide clinical governance with systems in clinical areas less well, with a mean value of 4.19.

Shortfall

The areas showing highest shortfall were the underpinning of clinical governance with systems in clinical areas, and a nominated clinical lead in clinical areas.

Figure 8.2.7: Corporate accountability items: importance / achievement / shortfall



Corporate accountability items	Means		
	Achievement	Importance	Shortfall
32 There is an executive director with responsibility for clinical governance	5.52	5.54	0.02
34 Clinical areas have a nominated clinical governance lead	4.89	5.35	0.46
35 There is a formal clinical governance committee, reporting to the board	5.70	5.59	-0.11
36 There are local arrangements to collate information for the CG committee	4.95	5.32	0.37
37 Organisation-wide CG is underpinned by systems in clinical areas	4.19	5.31	1.11

Table 8.2.7 Corporate accountability achievement data

Frequency tables for corporate accountability items						
	1.00	2.00	3.00	4.00	5.00	6.00
	Count	Count	Count	Count	Count	Count
A32	14	14	23	68	235	823
A34	34	52	91	170	310	520
A35	2	7	21	38	173	936
A36	9	38	85	210	367	468
A37	26	74	210	381	309	177

Frequency tables for corporate accountability items (%ages)						
	1.00	2.00	3.00	4.00	5.00	6.00
	%	%	%	%	%	%
A32	1.2%	1.2%	2.0%	5.8%	20.0%	69.9%
A34	2.9%	4.4%	7.7%	14.4%	26.3%	44.2%
A35	.2%	.6%	1.8%	3.2%	14.7%	79.5%
A36	.8%	3.2%	7.2%	17.8%	31.2%	39.8%
A37	2.2%	6.3%	17.8%	32.4%	26.3%	15.0%

Leadership and collaboration

The nineteen items relating to leadership and collaboration concern leadership, building clinical teams, involving service users, collaboration with local health and social care partners and organisational learning.

Importance

Again, there was limited variation between items, with most scoring over five. Items 51 and 52 concerning local health and social care collaboration were rated slightly lower than others were, as was item 42 on sharing clinical protocols across agencies.

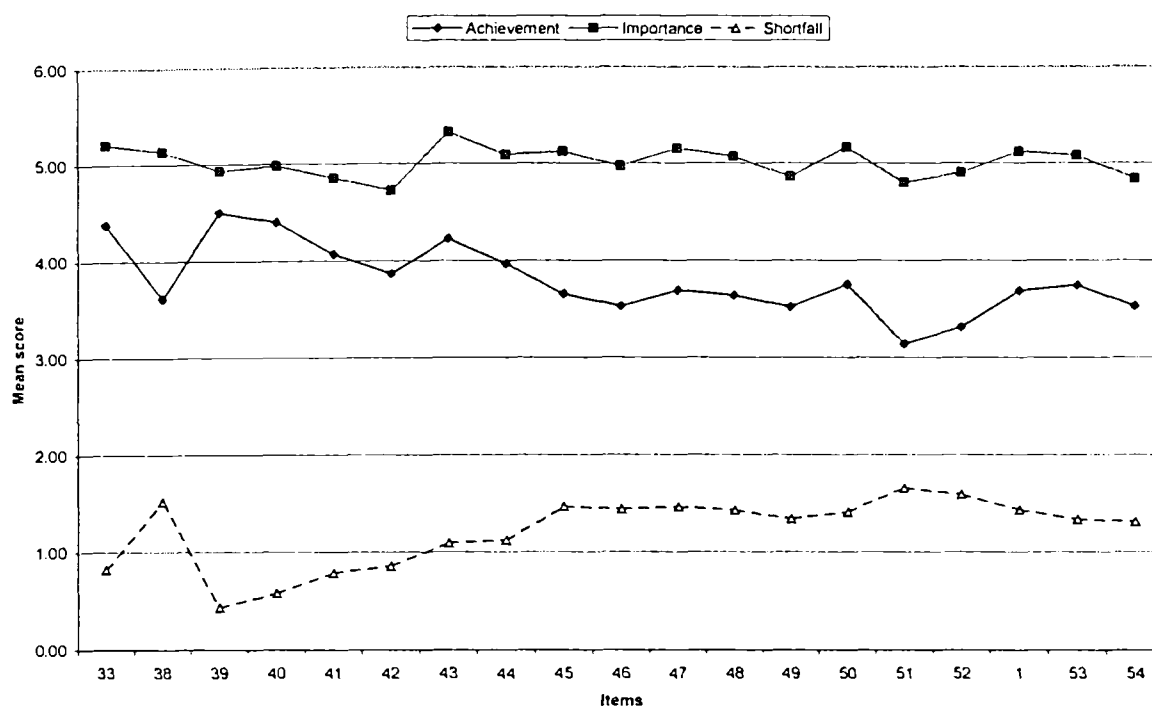
Achievement

These items were generally less well scored than the others, showing only moderate achievement. Items 51 and 52, concerning joint clinical governance arrangements and partnerships with health and social care agencies, scored particularly low. Items concerning building clinical teams and user involvement achieved a mean score of around 3.5, indicating modest achievement.

Shortfall

Given modest levels of achievement, there was considerable shortfall across many items. Particularly notable are those items relating to collaboration with local partners, user involvement and clinical teams.

Figure 8.2.8: Leadership & collaboration items: achievement / importance / shortfall



Leadership & collaboration items	Means		
	Achievement	Importance	Shortfall
33 Service delivery plans include quality improvement activity	4.38	5.21	0.82
38 Decisions about service developments / cutbacks are made on clear criteria	3.62	5.14	1.52
39 Local & national priorities are used to prioritise service development	4.50	4.94	0.43
40 NSFs are integrated with business planning / quality improvement activity	4.40	4.99	0.58
41 External guidelines are critically appraised before local adoption	4.07	4.86	0.79
42 Clinical protocols are shared with staff who work outside this organisation	3.87	4.73	0.86
43 The organisation shares a common vision for clinical governance	4.23	5.32	1.10
44 Leadership skills are identified and developed through training programmes	3.97	5.09	1.12
45 Clinical teams receive performance feedback	3.66	5.12	1.47
46 Clinical teams respond to changes by reorganising work processes	3.53	4.97	1.44
47 Staff have clear and shared objectives	3.69	5.15	1.46
48 There are clear processes for involving service users in service development	3.64	5.07	1.43
49 There are clear criteria for establishing user involvement groups	3.52	4.87	1.34
50 Service improvement activity focuses on the patient experience of care	3.75	5.17	1.41
51 Local health & social care agencies work jointly on clinical governance	3.14	4.80	1.66
52 Partnerships with health & social care agencies have clear, shared purposes	3.31	4.91	1.59
1 Teams work across boundaries to address quality issues	3.69	5.12	1.44
53 Staff are able to develop new and innovative models of service	3.75	5.08	1.34
54 Staff evaluate the best ways of training and learning from experience	3.53	4.84	1.31

Table 8.2.8: Leadership & collaboration achievement data

Frequency tables for leadership & collaboration items						
	1.00	2.00	3.00	4.00	5.00	6.00
	Count	Count	Count	Count	Count	Count
A33	13	63	158	362	387	194
A38	51	177	284	385	218	62
A39	13	48	130	328	460	198
A40	13	50	176	335	415	188
A41	19	86	233	397	339	103
A42	26	139	282	348	279	103
A43	36	83	176	345	357	180
A44	55	128	217	335	283	159
A45	47	153	323	361	220	73
A46	38	155	368	410	172	34
A47	35	120	319	444	217	42
A48	46	194	280	356	220	81
A49	62	210	312	310	208	75
A50	28	145	284	422	227	71
A51	110	282	345	265	118	57
A52	92	234	336	302	153	60
A1	11	124	341	488	174	39
A53	25	116	318	438	235	45
A54	24	162	384	410	167	30

Frequency tables for leadership & collaboration items (%ages)						
	1.00	2.00	3.00	4.00	5.00	6.00
	%	%	%	%	%	%
A33	1.1%	5.4%	13.4%	30.8%	32.9%	16.5%
A38	4.3%		24.1%	32.7%	18.5%	5.3%
A39	1.1%	4.1%	11.0%	27.9%	39.1%	16.8%
A40	1.1%	4.2%	15.0%	28.5%	35.3%	16.0%
A41	1.6%	7.3%	19.8%	33.7%	28.8%	8.8%
A42	2.2%	11.8%	24.0%	29.6%	23.7%	8.8%
A43	3.1%	7.1%	15.0%	29.3%	30.3%	15.3%
A44	4.7%	10.9%	18.4%	28.5%	24.0%	13.5%
A45	4.0%	13.0%	27.4%	30.7%	18.7%	6.2%
A46	3.2%	13.2%	31.3%	34.8%	14.6%	2.9%
A47	3.0%	10.2%	27.1%	37.7%	18.4%	3.6%
A48	3.9%	16.5%	23.8%	30.2%	18.7%	6.9%
A49	5.3%	17.8%	26.5%	26.3%	17.7%	6.4%
A50	2.4%	12.3%	24.1%	35.9%	19.3%	6.0%
A51	9.3%	24.0%	29.3%	22.5%	10.0%	4.8%
A52	7.8%	19.9%	28.5%	25.7%	13.0%	5.1%
A1	.9%	10.5%	29.0%	41.5%	14.8%	3.3%
A53	2.1%	9.9%	27.0%	37.2%	20.0%	3.8%
A54	2.0%	13.8%	32.6%	34.8%	14.2%	2.5%

3. Which domains show most achievement, and where are the shortfalls?

In order to identify progress between OPCG domains, mean scores for items were aggregated; results are summarised in figure 8.2.9 below. As five OPCG domains contained different numbers of items, scores were standardised to between zero (low) and ten (high) for each domain, using the formula: $\text{Score} = (\text{sum of all domain items} - \text{no. of items}) / (5 * \text{no. of items}) * 10$

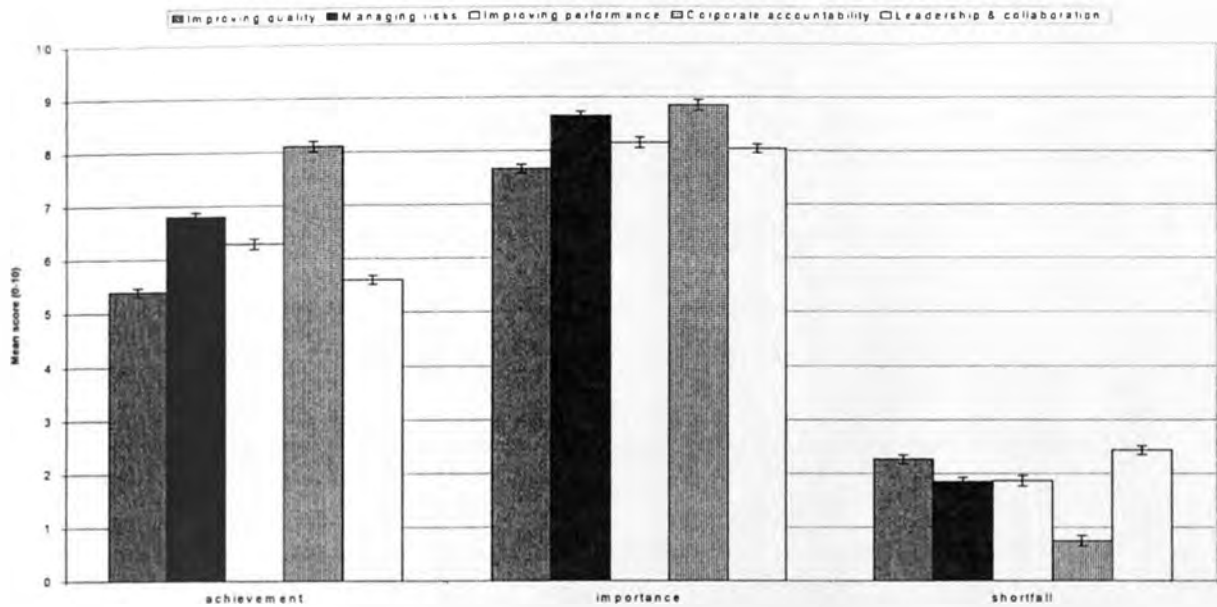
Achievement

Mean achievement scores ranged from 8.10 to 5.40, indicating room for improvement across all. The domains showing highest achievement across the sample were corporate accountability and risk management, respectively scoring 8.10 (95% CI 8.28 – 7.90) and 6.80 (95% CI 6.98 – 6.62). This is encouraging, as it shows that the priorities of corporate accountability and risk management were being addressed. However, there was evidence of less achievement in quality improvement and leadership and collaboration, with mean scores of 5.40 (95% CI 5.59 – 5.21) and 5.60 (95% CI 5.79 – 5.41) respectively.

Shortfall

Corporate accountability shows the lowest mean shortfall of .74 (95% CI .82 - .66), evidence of good achievement in this important domain. Risk management and performance improvement had a mean shortfall of 1.84 (95% CI 1.92 – 1.76) and 1.86 (95% CI 1.96 – 1.76) respectively, showing the need for continued work. The largest shortfalls were in the areas of quality improvement and leadership & collaboration, with mean shortfalls of 2.26 (95% CI 2.34 – 2.28) and 2.43 (95% CI 2.51 – 2.35) respectively.

Figure 8.2.9: Perceived achievement, importance and shortfall across domains



Importance scales (0-10)

	Count	Minimum	Maximum	Mean	Standard Error of Mean	Std Deviation
Improving quality	1177	1.82	10.00	7.66	.04	1.21
Managing risks	1177	.77	10.00	8.65	.03	1.13
Performance improvement	1177	1.33	10.00	8.15	.04	1.41
Corporate accountability	1177	.40	10.00	8.85	.04	1.30
Leadership & collaboration	1177	.63	10.00	8.04	.04	1.26

Achievement scales (0-10)

	Count	Minimum	Maximum	Mean	Standard Error of Mean	Std Deviation
Improving quality	1177	.36	9.45	5.40	.04	1.44
Managing risks	1177	.31	10.00	6.81	.04	1.51
Improving performance	1177	.00	10.00	6.29	.05	1.88
Corporate accountability	1177	1.60	10.00	8.10	.05	1.56
Leadership & collaboration	1177	.32	9.68	5.61	.04	1.52

Shortfall (0-10)

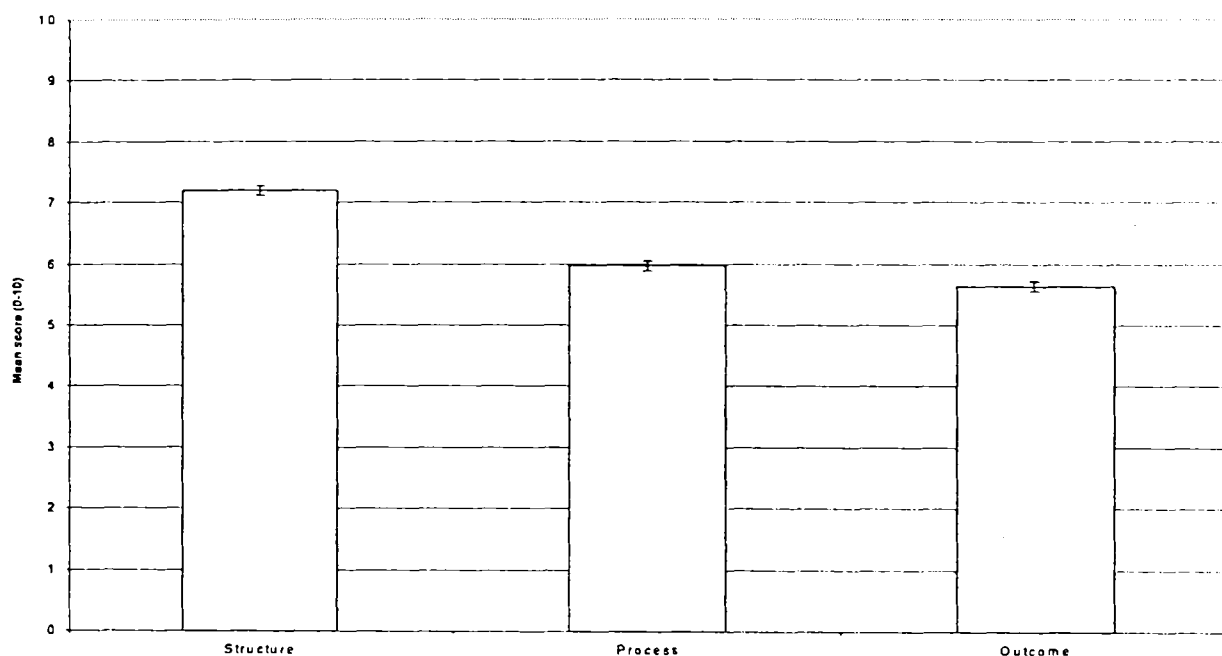
	Count	Minimum	Maximum	Mean	Standard Error of Mean	Std Deviation
Improving quality	1177	-4.55	8.00	2.26	.04	1.48
Managing risks	1177	-2.77	7.08	1.84	.04	1.38
Performance improvement	1177	-3.67	9.00	1.86	.05	1.77
Corporate accountability	1177	-6.80	8.40	.74	.04	1.42
Leadership & collaboration	1177	-1.47	8.32	2.43	.04	1.48

4. What progress is there in clinical governance structures, processes and outcomes?

Item scores were further aggregated under the heading 'structure', 'process' or 'outcome' and standardised between 0-10 to aid comparison (Figure 8.2.9). Details of the competency items assigned under each heading are in appendix 13, and the formula for standardisation was the same one used to aggregate scores by domain:

$$\text{Score} = (\text{sum of all heading items} - \text{no. of items}) / (5 * \text{no. of items}) * 10$$

Figure 8.2.10: Perceived achievement against structure, process and outcome



Structure, process and outcome achievement (0-10)

	Count	Minimum	Maximum	Mean	SE of mean	Std Deviation
Structure	1177	1.45	10.00	7.20	.04	1.36
Process	1177	.44	9.81	5.98	.04	1.37
Outcome	1177	.20	9.60	5.65	.04	1.41

Perhaps unsurprisingly, respondents perceived that most progress had been made against items relating to structural change (e.g. 35), and there had been rather less progress against process (e.g. 48) and outcome (e.g. 53). This is consistent with the perceived importance of the corporate accountability agenda identified above. While good progress has been made against the structural agenda, there is some considerable way to go in order to realise the long-term outcomes claimed for clinical governance.

5. How much variation in achievement is there between trusts?

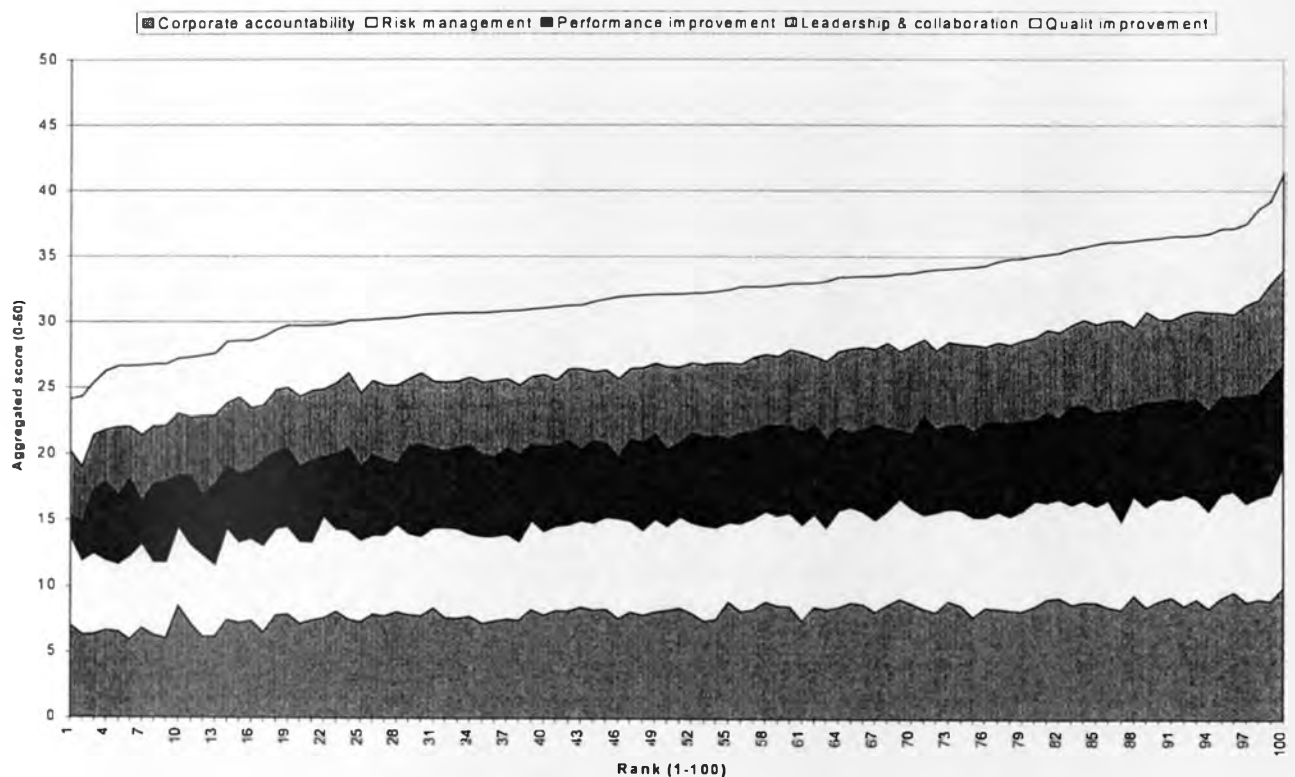
To facilitate comparisons between trusts, achievement scores were aggregated to produce trust means. The move from individual to organisational level analysis is only valid if there is more variation in scores of individuals *between* different trusts than there is in the scores of individuals *within* trusts. A simple one-way (non-repeated measure) analysis of variance tested this assumption, and results showed that the condition was satisfied for all domains (Table 8.2.9).

Table 8.2.9: Analysis of variance: achievement scores by trust (1-100)

ANOVA: Achievement by trust						
		Sum of Squares	df	Mean Square	F	Sig.
Improving quality	Between Groups	473.791	99	4.786	2.604	.000
	Within Groups	1979.674	1077	1.838		
	Total	2453.465	1176			
Managing risks	Between Groups	520.994	99	5.263	2.606	.000
	Within Groups	2175.292	1077	2.020		
	Total	2696.286	1176			
Improving performance	Between Groups	1190.277	99	12.023	4.340	.000
	Within Groups	2983.481	1077	2.770		
	Total	4173.758	1176			
Corporate accountability	Between Groups	794.347	99	8.024	4.196	.000
	Within Groups	2059.495	1077	1.912		
	Total	2853.842	1176			
Leadership & collaboration	Between Groups	635.322	99	6.417	3.340	.000
	Within Groups	2069.364	1077	1.921		
	Total	2704.686	1176			

As scores varied more *between* trusts than *within* them, domain scores were aggregated by trust (0-10), summed to produce summary trust scores (0-50), and ranked from lowest to highest (Figure 8.2.11). Trusts with the lowest global scores tended to score particularly poorly on quality improvement, leadership and collaboration and performance improvement domains. Across the sample, corporate accountability scored quite highly, with quality improvement and leadership and collaboration showing rather more room for improvement.

Figure 8.2.11: Ranked aggregate domain scores by trust



Ranked aggregated mean scores on the five OPGC domains across all 100 trusts in the study.

6. Multivariate analysis of achievement on OPCG domains, structure / process/ outcome aggregates, and high / medium / low competency aggregates by trust type, respondent and CHI visit- status

Sample respondents differed on the type of trust in which they worked (acute, ambulance or MH/LD), their position in the organisation (board member or directorate manager), and the CHI visit-status of the trust in which they worked (not yet given a date for first visit, undergoing the process, having received a report 3+ months ago). In order to assess the effect of each of these independent variables on OPCG domain scores, structure/ process/ outcome aggregates and low / medium / high difficulty aggregates, three separate Multiple Analysis of Variance (MANOVA) analyses were undertaken. In each case the full factorial model was assessed, including all interactions between the independent variables.

Additional data considerations

(a) Normal distribution of dependent variables

While significance tests for multivariate techniques are based the assumption that the sampling distributions of means of dependent variables are normally distributed, with large samples the central limit theorem suggests that sampling distributions of means approaches normality even when raw scores do not (Tabachnick and Fidell 2001). Given the large sample size, multivariate normality is not an issue within the present dataset.

(b) Screening for multivariate outliers

Cases were screened for multivariate outliers using SPSS REGRESSION, specifying RESIDUALS=OUTLIERS MAHAL, with case id number as the dependent variable. As no multivariate outliers were found, all 1177 cases were retained.

(c) Homogeneity of variance-covariance

Box's M tests the assumption that variance-covariance matrices within each cell are sampled from the same population variance-covariance matrix and can be pooled to create a single error estimate. Box's-M was statistically significant for both OPCG domain ($F = 1.737$; $d.f. = 33,329$; $p = .001$) and structure/ process/ outcome ($F = 1.546$; $d.f. = 38,639$; $p = .001$) analyses, which suggests that the assumption was violated and that results of these analyses should be considered tentative. For the level of difficulty aggregate analysis, Box's-M was not significant ($F = 1.323$, $d.f. = 38639$, $p = .030$), suggesting robust conclusions.

A: MANOVA of trust type, respondent status and CHI visit-status on OPCG domains

Main effects

Multivariate tests identified statistically significant main effects for trust type ($F = 10.391$, $d.f. = 10$, $P < .001$), CHI visit-status ($F = 4.440$, $d.f. = 10$, $P < .001$), board member status ($F = 4.553$, $d.f. = 5$, $P < .001$) and the interaction between trust type and CHI visit-status ($F = 5.346$, $d.f. = 10$, $P < .001$).

Table 8.2.10: Multivariate tests of significance (Pillai's trace and Wilkes' lambda)

Effect	Test	Value	d.f.	Error d.f.	F	Power	Sig.
TYPE	Pillai's Trace	.085	10.000	2320.000	10.337	1	.000**
	Wilkes' Lambda	.916	10.000	2318.000	10.391a	1	.000**
CHID	Pillai's Trace	.038	10.000	2320.000	4.440	.999	.000**
	Wilkes' Lambda	.963	10.000	2318.000	4.440a	.999	.000**
BOARDMEM	Pillai's Trace	.019	5.000	1159.000	4.553a	.973	.000**
	Wilkes' Lambda	.981	5.000	1159.000	4.553a	.973	.000**
TYPE * CHID	Pillai's Trace	.045	10.000	2320.000	5.299	1	.000**
	Wilkes' Lambda	.955	10.000	2318.000	5.346a	1	.000**
TYPE * BOARDMEM	Pillai's Trace	.012	10.000	2320.000	1.426	.731	.162
	Wilkes' Lambda	.988	10.000	2318.00	1.427a	.731	.162
CHID * BOARDMEM	Pillai's Trace	.004	10.000	2320.000	.450	.240	.922
	Wilkes' Lambda	.996	10.000	2318.000	.449a	.240	.922
TYPE * CHID * BOARDMEM	Pillai's Trace	.003	10.000	2320.000	.348	.188	.968
	Wilkes' Lambda	.997	10.000	2318.000	.348a	.188	.968

a= Exact statistic. ** = $p < .001$

Given significant multivariate effects, the nature of relationships between independent and dependent variables was explored further. However, while non-significant univariate tests between independent and individual dependent variables uncovered by this method are robust, significant results may be misleading when dependent variables are correlated (Tabachnick and Fidell 2001). Under such circumstances, there is a possibility of false significant effects due simply to the variance shared between correlated dependent variables. In order to assess the extent of the problem in the proposed analysis, Pearson correlations (r) were requested for the OPCG domain scores (Table 8.2.11).

Table 8.2.11: Correlations between dependent variables (Pearson's r)

	Improving quality	Managing risks	Improving performance	Corporate accountability
Managing risks	.672***	1		
Improving performance	.537***	.529***	1	
Corporate accountability	.557***	.597***	.477***	1
Leadership & collaboration	.686***	.689***	.590***	.622***

(n) = 1177. *** = $p < .001$ (2-tailed). Raw OPGC item scores summed and standardised to produce five factor scores of 0-10.

As correlations between domain scores were positive and in excess of .3, simple univariate F scores reported to clarify the effects between dependent and independent variables should be interpreted with caution. Two further courses of action were considered: either to statistically control for covariation in the dependent variables using Roy-Bargmann step-down F analysis; or compute uncorrelated regression factor scores for OPGC domains taken from the factor analysis detailed in section 6.2.1 above.

While the use of step-down F allowed retention of the simple summated domain scores, its use is problematic and poses severe difficulties in interpretation. The procedure requires ordering of entry of dependent variables, and while in the present case there was no compelling theoretical case for prioritising entry, entry could be ordered on the basis of univariate F scores, with most significant dependent variables entered first. More importantly, interpretation of step-down F becomes extremely difficult in circumstances where in the presence of higher-order dependent variables as covariates, a previously non-significant dependent variable suddenly becomes significant (Tabachnick and Fidell 2001). This proved to be the case in the present analysis, reported in appendix 14. Such a situation

does little to clarify the effects between dependent and independent variables, and so this course of action was rejected for final reporting.

In contrast to step-down F there are no technical difficulties associated with using uncorrelated regression factor scores; there is no longer any covariation between dependent variables. However, this course of action required the use of estimates of domain scores based on a prior factor analysis that were different to the simple summated domain scores used above. As simple summation was selected to enable data collected within the present study to be used subsequently as benchmarks, an important consideration given the intended use of the OPCG as a tool for tracking change over time, the replacement of simple summated scores was also rejected. The analysis reported is of univariate F-tests of significance for those areas with significant multivariate effects (table 8.2.12). Results should be interpreted with caution in the light of correlations between the dependent variables reported above.

Table 8.2.12: Tests of trust type, CHI visit and board membership, and interactions

IV	DV	Univariate F of IV on DV	d.f.	Sig.	Observed power (a)
TYPE	Improving quality	6.064**(b)	2	.002	.886
	Managing risks	.492	2	.612	.131
	Improving performance	9.148**	2	.000	.976
	Corporate accountability	7.521**(b)	2	.001	.944
	Leadership & collaboration	2.560	2	.078	.513
CHID	Improving quality	1.594	2	.203	.339
	Managing risks	.474	2	.622	.128
	Improving performance	7.374**	2	.001	.940
	Corporate accountability	3.80* (b)	2	.022	.696
	Leadership & collaboration	.793	2	.453	.186
BOARDMEM	Improving quality	6.661**(b)	1	.010	.732
	Managing risks	8.509**(b)	1	.004	.830
	Improving performance	.270	1	.603	.081
	Corporate accountability	16.351**	1	.000	.981
	Leadership & collaboration	10.122**(b)	1	.002	.889
TYPE * CHID	Improving quality	1.081	2	.340	.240
	Managing risks	4.654**(b)	2	.010	.784
	Improving performance	5.417**	2	.005	.846
	Corporate accountability	1.108	2	.331	.246
	Leadership & collaboration	.282	2	.754	.095

(a) Computed at alpha = .05

(b) Significant secondary effects should be treated with caution given correlations between dependent variables

** $p < .01$

* $p < .05$

(a) Trust type

Univariate analysis identified a significant effect of trust type on scores for improving performance ($F = 9.148$, $d.f.=2$, $P = .002$), corporate accountability ($F = 7.521$, $d.f.=2$, $P = .001$), and improving quality ($F = 6.064$, $d.f. = 2$, $P = .002$). Further post-hoc analysis (Scheffe) revealed that on improving performance, the overall effect was due to ambulance trusts scoring significantly lower than both acute ($P < .001$) and MH/LD/Social care ($P <$

.001) trusts. Acute trusts scored higher than ambulance trusts on corporate accountability ($P = .019$), and higher than MH/LD/Social care trusts on improving quality ($P = .003$).

Table 8.2.13: Between trust variation

Trust-type descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Improving performance	acute	837	6.4357	1.74452	.06030	6.3173	6.5540	.00	10.00
	ambulance	96	5.2500	2.60499	.26587	4.7222	5.7778	.00	10.00
	MH/LD/Social care	244	6.2199	1.88353	.12058	5.9824	6.4575	.00	10.00
	Total	1177	6.2943	1.88391	.05491	6.1865	6.4020	.00	10.00
Corporate accountability	acute	837	8.1711	1.52359	.05266	8.0677	8.2745	1.60	10.00
	ambulance	96	7.7042	1.84013	.18781	7.3313	8.0770	3.20	10.00
	MH/LD/Social care	244	8.0246	1.53276	.09812	7.8313	8.2179	1.60	10.00
	Total	1177	8.1026	1.55780	.04541	8.0135	8.1917	1.60	10.00
Improving quality	acute	837	5.4800	1.42838	.04937	5.3831	5.5769	.36	9.45
	ambulance	96	5.3845	1.63825	.16720	5.0525	5.7164	1.09	8.55
	MH/LD/Social care	244	5.1200	1.38826	.08887	4.9449	5.2950	1.45	8.36
	Total	1177	5.3975	1.44440	.04210	5.3149	5.4801	.36	9.45

(b) CHI visit-status

CHI visit-status significantly affected scores on improving performance ($F = 7.374$, d.f. =2, $P = .001$) and corporate accountability ($F = 3.800$, d.f. =2, $P = .022$). The latter significant effect should be regarded with caution; it reaches significance only at the .05 level, and as noted earlier, improving performance and corporate accountability scores were moderately correlated (Pearson's $r = .477$, $p < .001$). Post-hoc analysis (Scheffe) showed that for improving performance those yet to receive a first visit date scored lower than those undergoing the process ($P = .007$). On corporate accountability, those undergoing the visit process scored higher than both those yet to receive a first visit date ($P = .003$) and those who had received a report ($P = .033$). Scores on both dependent variables suggest a temporary heightening of activity, falling away once CHI reports.

Table 8.2.14: between CHI-visit status differences

CHI visit-status descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Improving performance	none	442	6.0830	2.01935	.09605	5.8942	6.2717	.00	10.00
	undergoing visit process	275	6.5285	1.79288	.10811	6.3156	6.7413	1.00	10.00
	published report 3+ mths	460	6.3572	1.78231	.08310	6.1939	6.5206	.00	10.00
	Total	1177	6.2943	1.88391	.05491	6.1865	6.4020	.00	10.00
Corporate accountability	none	442	7.9710	1.63096	.07758	7.8186	8.1235	1.60	10.00
	undergoing visit process	275	8.3738	1.47719	.08908	8.1985	8.5492	2.40	10.00
	published report 3+ mths	460	8.0670	1.51565	.07067	7.9281	8.2058	1.60	10.00
	Total	1177	8.1026	1.55780	.04541	8.0135	8.1917	1.60	10.00

(c) Boardmember status

Respondents were either board members or directorate-level managers. Respondent status significantly affected corporate accountability ($F = 16.351$, d.f. =2, $P < .001$), leadership and collaboration ($F = 10.122$, d.f. =2, $P = .002$), managing risks ($F = 8.509$, d.f. =2, $P = .004$), and improving quality ($F = 6.661$, d.f. =2, $P = .010$). Board-level respondents scored each of these domains higher than their directorate-level colleagues.

Table 8.2.15: Between board-member status differences

Boardmember descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Corporate accountability	Yes	555	8.2796	1.35324	.05744	8.1668	8.3925	3.20	10.00
	No	622	7.9447	1.70555	.06839	7.8104	8.0790	1.60	10.00
	Total	1177	8.1026	1.55780	.04541	8.0135	8.1917	1.60	10.00
Leadership & collaboration	Yes	555	5.7984	1.37434	.05834	5.6838	5.9130	.95	9.58
	No	622	5.4354	1.61497	.06475	5.3083	5.5626	.32	9.68
	Total	1177	5.6066	1.51654	.04420	5.5199	5.6933	.32	9.68
Managing risks	Yes	555	6.9672	1.33714	.05676	6.8557	7.0786	.92	10.00
	No	622	6.6713	1.64483	.06595	6.5418	6.8008	.31	10.00
	Total	1177	6.8108	1.51419	.04414	6.7242	6.8974	.31	10.00
Improving quality	Yes	555	5.5181	1.34947	.05728	5.4056	5.6306	1.09	9.27
	No	622	5.2900	1.51712	.06083	5.1705	5.4094	.36	9.45
	Total	1177	5.3975	1.44440	.04210	5.3149	5.4801	.36	9.45

(d) Interaction of trust type with CHI visit-status

There was a significant interaction effect between trust type and CHI visit-status on improving performance ($F = 5.417$, d.f. =2, $P = .005$) and managing risks ($F = 4.654$, d.f. =2, $P = .010$). Estimated marginal means are reproduced in table 8.2.16 below.

Table 8.2.16: Interaction effects

Trust type * CHI'd? descriptives						
Dependent Variable	Trust type	CHI'd?	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Improving performance	acute	none	6.446	.136	6.179	6.712
		undergoing visit process	6.590	.136	6.323	6.857
		published report 3+ mths	6.345	.086	6.176	6.514
	ambulance	none	4.655	.229	4.207	5.104
		undergoing visit process	6.343	.376	5.605	7.080
		published report 3+ mths	^a			
	MH/LD/Social care	none	6.187 ^b	.136	5.921	6.454
		undergoing visit process	6.347 ^b	.238	5.880	6.813
		published report 3+ mths	^a			
Managing risks	acute	none	6.854	.111	6.636	7.071
		undergoing visit process	6.940	.111	6.722	7.158
		published report 3+ mths	6.806	.070	6.668	6.944
	ambulance	none	7.232	.187	6.866	7.598
		undergoing visit process	6.444	.307	5.843	7.046
		published report 3+ mths	^a			
	MH/LD/Social care	none	6.512 ^b	.111	6.295	6.730
		undergoing visit process	7.011 ^b	.194	6.631	7.392
		published report 3+ mths	^a			

^a. This level combination of factors is not observed, thus the corresponding population marginal mean is not estimable.

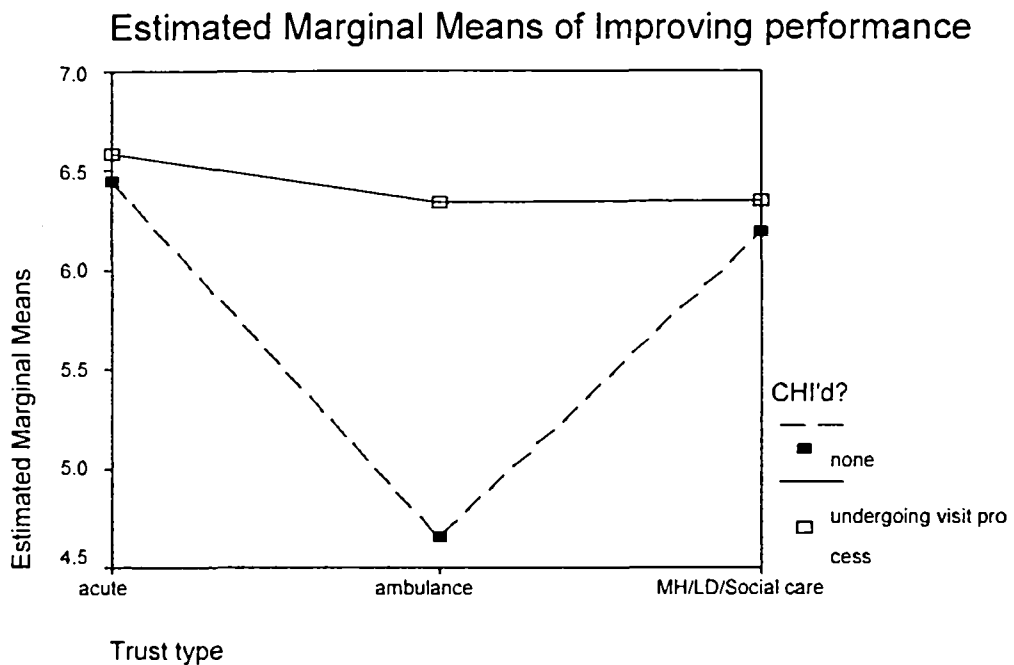
^b. Based on modified population marginal mean.

To further explore and clarify the interaction effects of trust type and CHI visit-status on these two dependent variables, estimated marginal means were plotted against each other under different levels of the independent variables of trust type and CHI visit-status for improving performance (fig 8.2.12) and managing risks (fig 8.2.13).

(i) Effect on performance improvement

There was a significant, positive effect of undergoing a CHI review process on performance improvement for ambulance trusts, but no significant effect for acute and MH/LD/Social care trusts (figure 8.2.12).

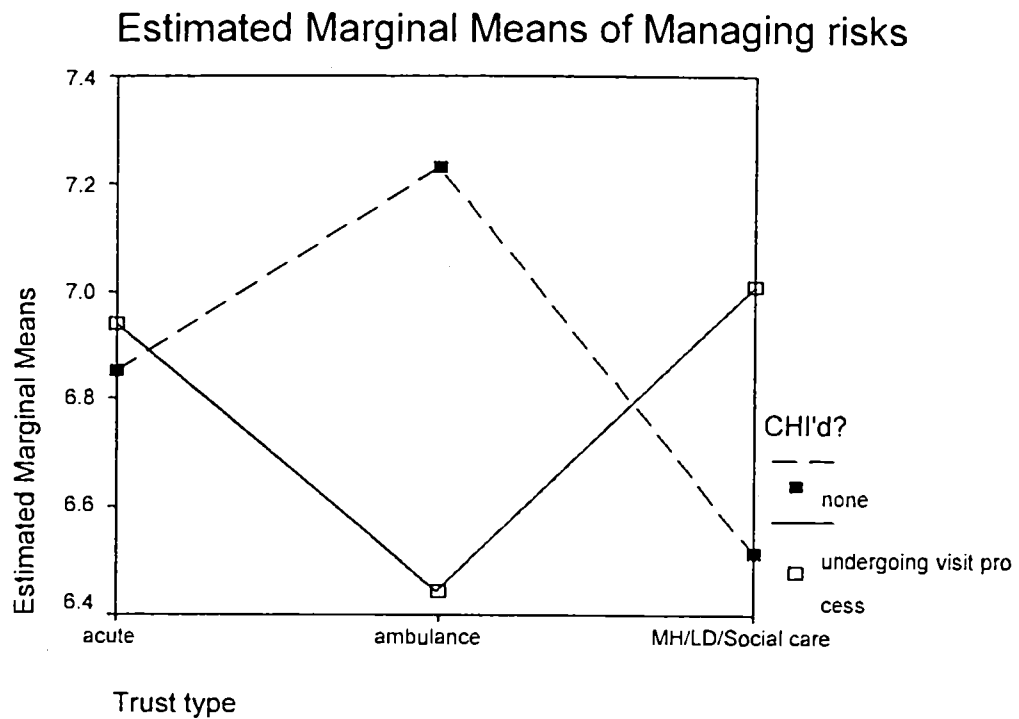
Figure 8.2.12: Interaction effects between trust type and CHI visit-status on improving performance



(ii) Effect on managing risks

There was a non-significant, negative effect of undergoing a CHI review on managing risks in ambulance trusts and a non-significant positive effect for acute and MH/LD/Social care trusts (figure 8.2.13).

Figure 8.2.13: Interaction effects between trust type and CHI visit-status on managing risks



B: MANOVA of trust type, respondent status and CHI visit-status on structure/ process/ outcome aggregates

Main effects

Multivariate tests identified statistically significant main effects for each of the three independent variables: trust type ($F = 25.782$, d.f. = 10, $p < .001$); respondent's level in the organisation ($F = 3.777$, d.f. = 3, $p = .010$); and CHI visit-status ($F = 3.150$, d.f. = 6, $p = .004$).

Table 8.2.17: Multivariate tests of significance (Pillai's trace and Wilkes' lambda)

Effect	Test	Value	d.f.	Error d.f.	F	Power	Sig.
TYPE	Pillai's Trace	.085	3.000	2324.000	25.782	1	.000**
	Wilkes' Lambda	.916	3.000	2322.000	25.839a	1	.000**
CHID	Pillai's Trace	.038	6.000	2324.000	3.150	.925	.004**
	Wilkes' Lambda	.963	6.000	2322.000	3.159a	.925	.004**
BOARDMEM	Pillai's Trace	.019	3.000	1161.000	3.777a	.973	.010**
	Wilkes' Lambda	.981	3.000	1161.000	3.777a	.973	.010**
TYPE * CHID	Pillai's Trace	.005	6.000	2324.000	.992	.399	.429
	Wilkes' Lambda	.955	6.000	2322.000	.993a	.399	.428
TYPE * BOARDMEM	Pillai's Trace	.016	6.000	2324.000	3.095	.919	.005**
	Wilkes' Lambda	.984	6.000	2322.000	3.100	.920	.005**
CHID * BOARDMEM	Pillai's Trace	.004	6000	2324.000	.755	.304	.605
	Wilkes' Lambda	.996	6.000	2322.000	.449a	.240	.605
TYPE * CHID * BOARDMEM	Pillai's Trace	.004	6000	2324.000	.845	.340	.535
	Wilkes' Lambda	.996	6000	2322.000	.845	.340	.535

a= Exact statistic. ** = $p < .001$

Given significant multivariate effects, the nature of relationships between independent and dependent variables was explored further. For consistency with the previous analysis, the analysis reported is of the univariate F-tests of significance for those areas with significant multivariate effects (table 8.2.18). In order to assess the extent of possible over-estimation of significance caused by correlation between dependent variables, Pearson correlations (r) for structure/ process/ outcome aggregates are given in table 8.2.19. Given the high correlations between the three scales, significant effects of second or third dependent variables should be treated with caution.

Table 8.2.19: Correlations between dependent variables (Pearson's r)

	Structure	Process
Process	.804***	1
Outcome	.707***	.869***

(n) = 1177. *** = $p < .001$ (2-tailed). Raw OPGC item scores summed and standardised to produce scores for structure, process and outcome 0-10.

Table 8.2.18: Tests of trust type, CHI visit and board membership, and their interactions

IV	DV	Univariate F of IV on DV	d.f.	Sig.	Observed power (a)
TYPE	Structure	12.279**	2	.000	.996
	Process	1.976	2	.139	.410
	Outcome	3.984* b	2	.019	.715
CHID	Structure	4.397*	2	.013	.759
	Process	.424	2	.655	.119
	Outcome	.761	2	.467	.180
BOARDMEM	Structure	11.050**	1	.001	.913
	Process	8.654**b	1	.003	.836
	Outcome	7.622**b	1	.006	.788
TYPE * BOARDMEM	Structure	1.515	2	.220	.324
	Process	.814	2	.443	.190
	Outcome	2.288	2	.102	.466

(a) Computed at alpha = .05

(b) Significant secondary effects should be treated with caution given correlations between dependent variables

** $p < .01$

* $p < .05$

(a) Trust type

Univariate analysis identified a significant effect of trust type on scores for structure ($F = 12.279$; $d.f.=2$; $P < .001$). Further post-hoc analysis (Scheffe) revealed that the overall effect was due to acute trusts scoring significantly higher than both ambulance ($P = .001$) and MH/LD/Social care ($P < .006$) trusts.

Table 8.2.20: Trust differences

Trust type descriptives								
Structure	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
acute	837	7.3086	1.33191	.04604	7.2182	7.3989	1.45	10.00
ambulance	96	6.7481	1.60531	.16384	6.4228	7.0734	2.73	9.27
MH/LD/Social care	244	6.9970	1.28676	.08238	6.8348	7.1593	3.09	9.64
Total	1177	7.1983	1.35831	.03959	7.1206	7.2759	1.45	10.00

(b) CHI visit-status

CHI visit-status significantly affected scores on structure at the 0.05 level ($F = 4.397$, d.f. = 2, $P = .013$). Post-hoc analysis (Scheffe) showed that this effect was due to those undergoing visits scoring higher than those who had yet to be given a visit date ($p = .007$). There was no significant difference between those undergoing the visit process and those who received a CHI visit report more than 3 months previously ($p = .170$), or those yet to be given a visit date and those who received a CHI visit report more than 3 months previously ($p = .331$). This suggests a small positive effect of CHI visits on structural items, limited to the time that the reviews are undertaken. There was no statistically significant effect on composite scores for processes or outcomes.

Table 8.2.21: CHI visit-status differences

CHI visit-status descriptives								
Structure	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
none	442	7.0708	1.39191	.06621	6.9406	7.2009	2.55	9.82
undergoing visit process	275	7.3950	1.26372	.07620	7.2450	7.5451	3.09	9.64
published report 3+ mths	460	7.2032	1.36875	.06382	7.0778	7.3286	1.45	10.00
Total	1177	7.1983	1.35831	.03959	7.1206	7.2759	1.45	10.00

(c) Boardmember status

Respondents were either board members or directorate-level managers. Respondent status significantly affected scores on structure ($F = 11.050$, $d.f. = 1$, $P = .001$), process ($F = 8.654$, $d.f. = 1$, $P = .003$), and outcome ($F = 7.622$, $d.f. = 1$, $P = .006$). On all three measures, board members scored achievement higher than their directorate-level colleagues did. These latter two effects should be treated with caution however; as noted earlier, scores for structure correlate highly with those for process (Pearson's $r = .804$, $p < .001$) and outcome (Pearson's $r = .707$, $p < .001$).

Table 8.2.22: Boardmember status differences

Board membership descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Structure	Yes	555	7.3284	1.21859	.05173	7.2268	7.4300	3.27	10.00
	No	622	7.0821	1.46306	.05866	6.9669	7.1973	1.45	9.82
	Total	1177	7.1983	1.35831	.03959	7.1206	7.2759	1.45	10.00
Process	Yes	555	6.1062	1.23015	.05222	6.0036	6.2088	2.06	9.81
	No	622	5.8656	1.47240	.05904	5.7496	5.9815	.44	9.81
	Total	1177	5.9790	1.36827	.03988	5.9008	6.0573	.44	9.81
Outcome	Yes	555	5.7769	1.31640	.05588	5.6672	5.8867	1.80	9.20
	No	622	5.5305	1.53543	.06157	5.4096	5.6514	.20	9.60
	Total	1177	5.6467	1.44098	.04200	5.5643	5.7291	.20	9.60

(b) Interaction of trust type with board membership status

In the initial multivariate analysis, there was a significant interaction effect between trust type and board membership status ($F = 3.095$, $d.f. = 6$, $p = .005$). However, univariate analysis, taking the effect of each dependent variable separately and reported in table 8.2.18

above, found no significant effect on any single dependent variable. This is due to the greater statistical power of multivariate analysis, in which the averaging of effects a variable across the levels of another variable reduces the number of subjects needed for a given degree of statistical power

Estimated marginal mean scores for dependent variables at different levels of trust type and board membership are reproduced in table 8.2.23, and relationships were further explored graphically by plotting the effects of the two independent variables on structure (fig. 8.2.14), process (fig. 8.2.15) and outcome (fig 8.2.16).

Table 8.2.23: Estimated marginal means

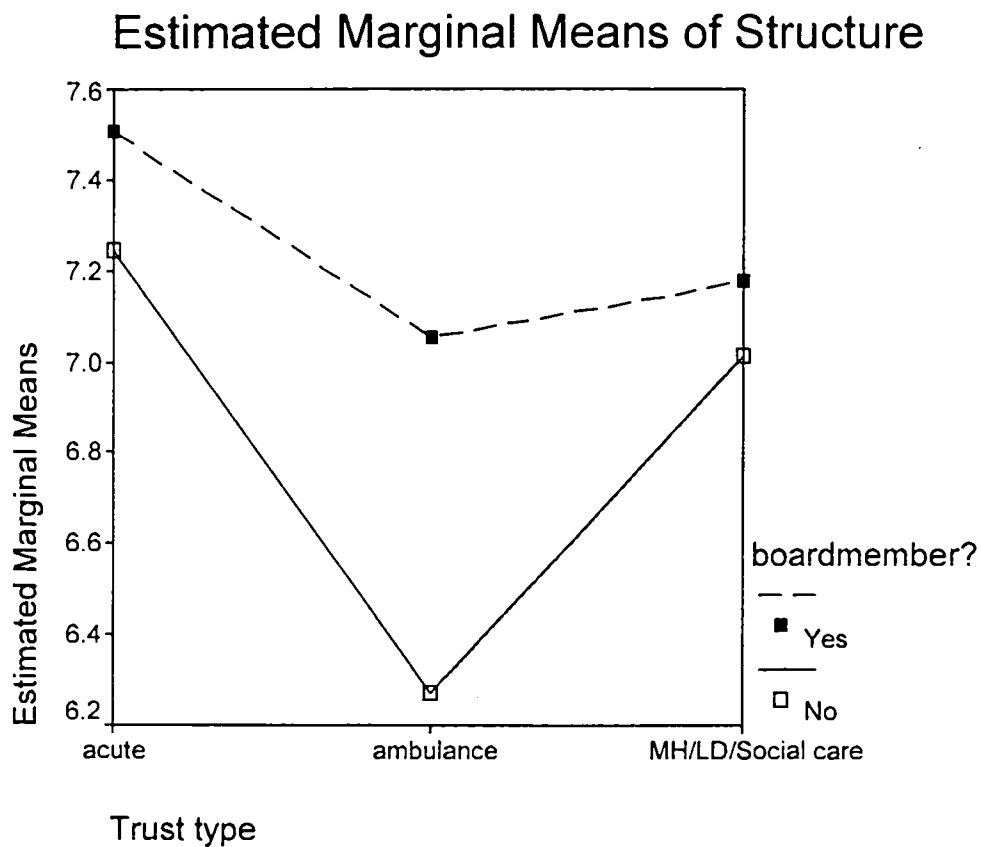
boardmember? * Trust type descriptives						
Dependent Variable	boardmember?	Trust type	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Structure	Yes	acute	7.505	.077	7.353	7.657
		ambulance	7.055 ^a	.188	6.687	7.423
		MH/LD/Social care	7.176 ^a	.138	6.905	7.447
	No	acute	7.247 ^a	.066	7.117	7.377
		ambulance	6.271 ^a	.258	5.764	6.778
		MH/LD/Social care	7.015 ^a	.143	6.735	7.295
Process	Yes	acute	6.082	.079	5.927	6.238
		ambulance	6.063 ^a	.192	5.686	6.439
		MH/LD/Social care	6.292 ^a	.141	6.015	6.569
	No	acute	5.891 ^a	.068	5.758	6.024
		ambulance	5.452 ^a	.264	4.933	5.970
		MH/LD/Social care	5.979 ^a	.146	5.693	6.266
Outcome	Yes	acute	5.838	.083	5.675	6.000
		ambulance	6.325 ^a	.201	5.931	6.719
		MH/LD/Social care	5.428 ^a	.148	5.138	5.719
	No	acute	5.589 ^a	.071	5.450	5.728
		ambulance	5.456 ^a	.277	4.913	5.998
		MH/LD/Social care	5.419 ^a	.153	5.119	5.719

a. Based on modified population marginal mean.

(i) Effect of trust type and boardmember status on structure

Respondents with board level positions scored achievement against structural items more highly than their directorate level colleagues in all trust types. The effect was strongest in ambulance trusts, although differences in effect between trust types did not reach statistical significance at the .05 level.

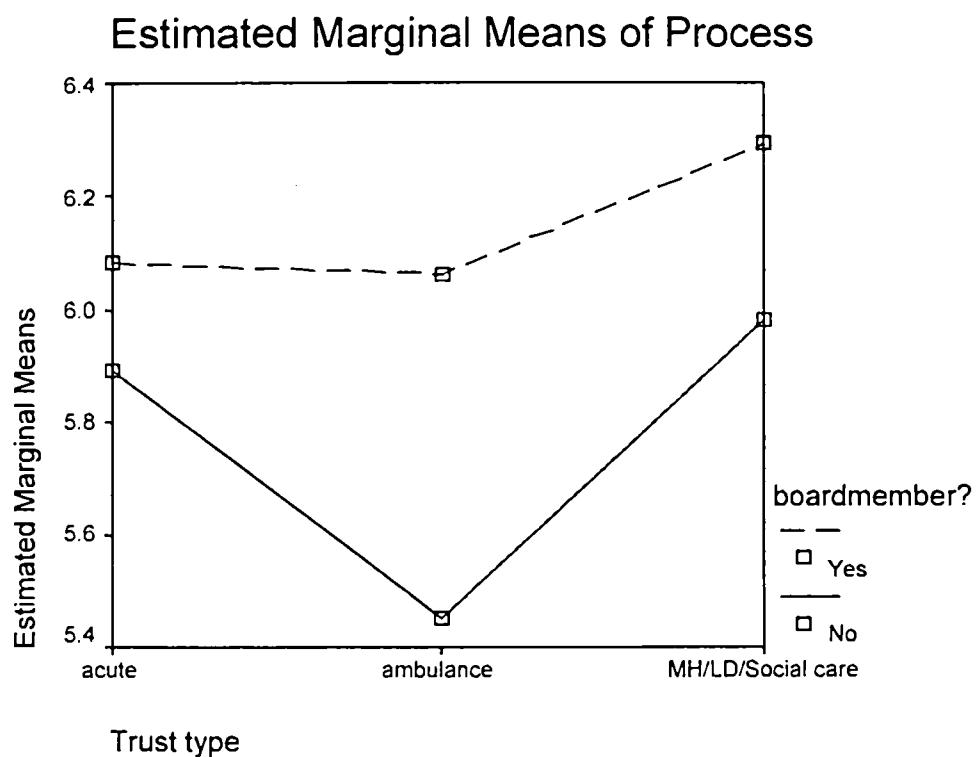
Figure 8.2.14: Marginal means of structure



(ii) Effect of trust type and boardmember status on process

As above with aggregates of structural items, respondents with board level positions scored achievement against process items more highly than their directorate level colleagues in all trust types. Again, although the effect was strongest within ambulance trusts, differences between trust types did not reach statistical significance at the .05 level.

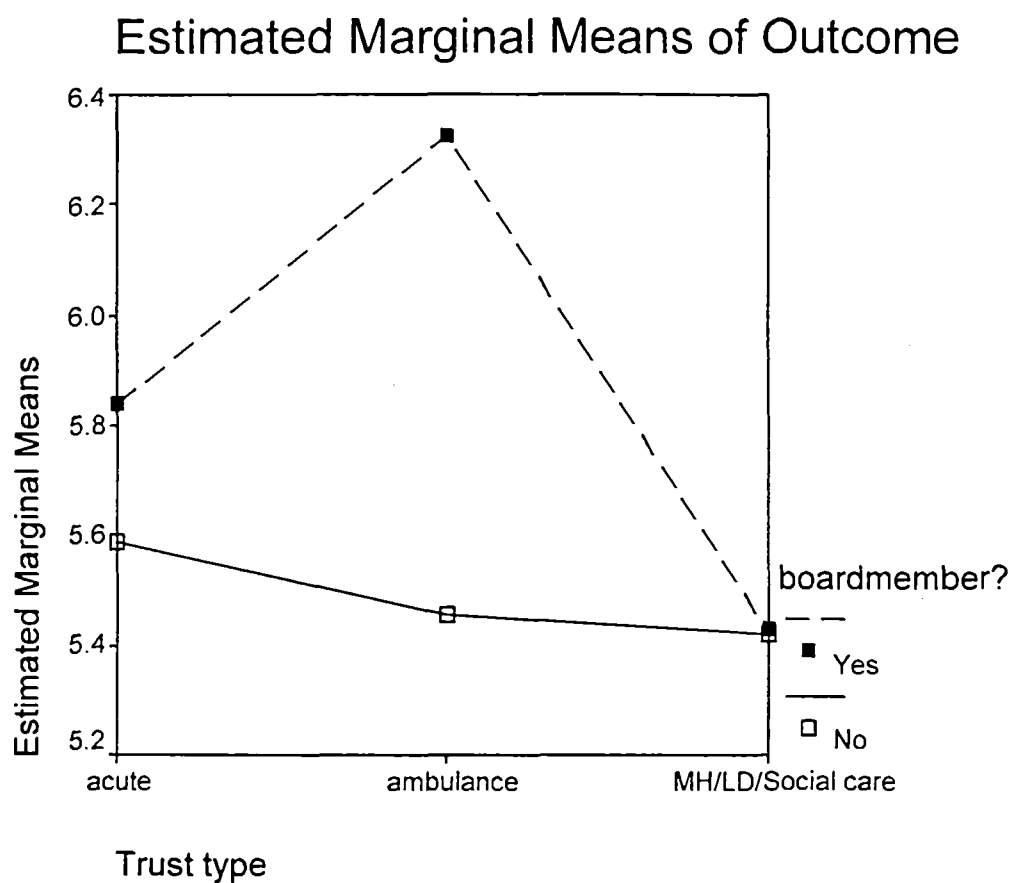
Figure 8.2.15: Marginal means of process



(iii) *Effect of trust type and boardmember status on outcome*

Although differences were not significant at the .05 level, respondents with board level positions in acute and ambulance trusts scored achievement against outcome items higher than their directorate level colleagues. This was not the case in MH/LD/Social care trusts, where there was no difference at all in scores.

Figure 8.2.16: Marginal means of outcome



C: MANOVA of trust type, respondent status and CHI visit-status on low, medium and high aggregates

Multivariate tests identified statistically significant main effects for trust type ($F = 12.431$, $d.f. = 6$, $P < .001$), CHI visit-status ($F = 5.252$, $d.f. = 6$, $P < .001$), and board member status ($F = 4.100$, $d.f. = 3$, $P < .007$), with no interaction effects significant at the .01 level.

Table 8.2.24: Multivariate tests of significance (Pillai's trace and Wilkes' lambda)

Effect	Test	Value	d.f.	Error d.f.	F	Power	Sig.
TYPE	Pillai's Trace	.062	6.000	2324.000	12.431	1	.000**
	Wilkes' Lambda	.938	6.000	2322.000	12.512a	1	.000**
CHID	Pillai's Trace	.027	6.000	2324.000	5.252	.996	.000**
	Wilkes' Lambda	.973	6.000	2322.000	5.271a	.996	.000**
BOARDMEM	Pillai's Trace	.010	3.000	1161.000	4.100a	.849	.007**
	Wilkes' Lambda	.990	3.000	1161.000	4.100a	.849	.007**
TYPE * CHID	Pillai's Trace	.012	6.000	2324.000	2.268	.798	.035
	Wilkes' Lambda	.988	6.000	2322.000	2.268a	.798	.035
TYPE * BOARDMEM	Pillai's Trace	.007	6.000	2324.000	1.432	.565	.198
	Wilkes' Lambda	.993	6.000	2322.000	1.433a	.565	.198
CHID * BOARDMEM	Pillai's Trace	.003	6.000	2324.000	.571	.232	.754
	Wilkes' Lambda	.997	6.000	2322.000	.571a	.232	.754
TYPE * CHID * BOARDMEM	Pillai's Trace	.002	6.000	2324.000	.348	.151	.911
	Wilkes' Lambda	.998	6.000	2322.000	.348a	.151	.911

a= Exact statistic. ** = $p < .001$

Where multivariate effects were statistically significant, the nature of relationships between independent and dependent variables was explored further using univariate F-tests (table 8.2.26). The extent of possible over-estimation of significance caused by correlation between dependent variables may be estimated from table 8.2.25. Given the high correlations between these scales, significant effects of second or third dependent variables should be treated with caution.

Table 8.2.25: Correlations between dependent variables (Pearson's r)

	Low	Medium
Medium	.839***	1
High	.774***	.857***

(n) = 1177. *** = $p < .001$ (2-tailed). Raw OPCG item scores summed and standardised to produce scores for low, medium and high difficulty (0-10).

Table 8.2.26: Tests of trust type, CHI visit and board membership

IV	DV	Univariate F of IV on DV	d.f.	Sig.	Observed power (a)
TYPE	Low	5.138**	2	.006	.825
	Medium	3.031*b	2	.049	.588
	High	1.273	2	.281	.277
CHID	Low	2.074	2	.126	.428
	Medium	2.899	2	.055	.567
	High	.160	2	.852	.075
BOARDMEM	Low	16.640**	1	.002	.885
	Medium	23.103***b	1	.001	.918
	High	13.755***b	1	.010	.735

(a) Computed at alpha = .05

(b) Significant secondary effects should be treated with caution given correlations between dependent variables

** $p < .01$

* $p < .05$

(a) Trust type

Univariate analysis identified a significant effect of trust type on scores for low ($F = 5.138$, $d.f. = 2$, $p = .006$) and medium ($F = 3.013$, $d.f. = 2$, $p = .049$) difficulty aggregates. Further post-hoc analysis (Scheffe) revealed the overall effect to be due to respondents from acute trusts scoring higher than those from MH/LD/Social care trusts on the medium difficulty aggregate ($p = .045$).

Table 8.2.27: Trust differences

Trust type descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
LO	acute	837	7.0903	1.28688	.04448	7.0030	7.1776	1.11	9.89
	ambulance	96	6.7882	1.45391	.14839	6.4936	7.0828	2.44	9.33
	MH/LD/Social care	244	6.9627	1.26422	.08093	6.8032	7.1221	3.00	9.56
	Total	1177	7.0392	1.29865	.03785	6.9649	7.1134	1.11	9.89
MED	acute	837	5.8086	1.42278	.04918	5.7120	5.9051	.24	9.65
	ambulance	96	5.6176	1.74148	.17774	5.2648	5.9705	.94	8.59
	MH/LD/Social care	244	5.5487	1.36093	.08712	5.3771	5.7203	2.00	8.35
	Total	1177	5.7391	1.44184	.04203	5.6567	5.8216	.24	9.65
HI	acute	837	5.6533	1.44602	.04998	5.5552	5.7514	.42	9.79
	ambulance	96	5.8070	1.47919	.15097	5.5073	6.1067	2.00	8.84
	MH/LD/Social care	244	5.8555	1.37325	.08791	5.6823	6.0286	1.16	8.95
	Total	1177	5.7077	1.43538	.04184	5.6256	5.7898	.42	9.79

(b) CHI visit status

CHI visit-status had a significant multivariate effect ($F = 5.252$, d.f. = 6, $p < .001$). While univariate analysis found no single dependent variable significant at the .05 level, post-hoc analysis (Scheffe) found that trusts undergoing the CHI review process scored higher on the medium difficulty aggregate than those yet to receive a visit date ($p = .014$).

Table 8.2.28: CHI visit-status differences

CHI visit -status multiple comparisons

Scheffe

Dependent Variable	(I) CHI'd?	(J) CHI'd?	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LO	none	none					
		undergoing visit process	-.2148	.09911	.096	-.4577	.0281
		published report 3+ mths	-.0458	.08595	.868	-.2565	.1648
	undergoing visit process	none	.2148	.09911	.096	-.0281	.4577
		undergoing visit process					
		published report 3+ mths	.1690	.09837	.229	-.0721	.4100
	published report 3+ mths	none	.0458	.08595	.868	-.1648	.2565
		undergoing visit process	-.1690	.09837	.229	-.4100	.0721
		published report 3+ mths					
MED	none	none					
		undergoing visit process	-.3227*	.10999	.014	-.5923	-.0531
		published report 3+ mths	-.1360	.09539	.362	-.3697	.0978
	undergoing visit process	none	.3227*	.10999	.014	.0531	.5923
		undergoing visit process					
		published report 3+ mths	.1867	.10917	.232	-.0808	.4543
	published report 3+ mths	none	.1360	.09539	.362	-.0978	.3697
		undergoing visit process	-.1867	.10917	.232	-.4543	.0808
		published report 3+ mths					
HI	none	none					
		undergoing visit process	.0007	.11001	1.000	-.2689	.2703
		published report 3+ mths	.0430	.09540	.903	-.1908	.2768
	undergoing visit process	none	-.0007	.11001	1.000	-.2703	.2689
		undergoing visit process					
		published report 3+ mths	.0423	.10918	.928	-.2253	.3099
	published report 3+ mths	none	-.0430	.09540	.903	-.2768	.1908
		undergoing visit process	-.0423	.10918	.928	-.3099	.2253
		published report 3+ mths					

Based on observed means.

*. The mean difference is significant at the .05 level.

(c) Board membership

Board members scored significantly higher than directorate-level managers did on low ($F=9.992$, d.f. =2, $p = .002$), medium ($F= 11.264$, d.f. =2, $p = .001$) and high ($F= 6.705$, d.f. =2, $p = .010$) difficulty aggregates. These effects should be treated with caution however, as scores for medium difficulty correlate highly with those for low (Person's $r = .839$ $p < .001$) and high (Pearson's $r = .774$, $p < .001$) difficulty.

8.3 Discussion

Study limitations

The study was designed to compare managers' perceptions of progress in clinical governance against a range of clinical governance competencies using the OPGC schedule. While respondents were randomly drawn from a stratified sample of all acute, ambulance and MH/LD/SC trusts in England, the usefulness and generalisability of the findings are affected by a number of methodological limitations.

Firstly, the study is limited to the perceptions of managers at board and directorate level within organisations, omitting the views of clinicians other than those with board level appointments such as medical and nursing directors. This limitation is deliberate, as the OPGC was designed specifically to gather data on competencies relating to organisational achievement from multiple respondents at board and directorate manager level. Managers' perceptions of progress are generally optimistic than those of clinical staff (Singer *et. al.* 2003), and this is likely to be all the more so given managers' responsibility for the clinical governance agenda. Indeed, results of the study of 'lived experience' of clinical governance among front line staff using the CGCQ reported in chapter six found that across multiple organisations, managers consistently rated the climate related to risk management and absence of unjust blame more favourably than their medical colleagues. Managers'

perceptions reported here are thus likely to be more favourable than the perceptions of other staff groups.

Evaluations of the early implementation of clinical governance, reviewed in chapter 2, identified different emphases between quality assurance and quality improvement evident in the approaches taken in primary and secondary care. It would have been informative to explore these differences in terms of the perceptions of progress in primary and secondary sectors across the OPCG domains by including primary sector organisations within the sample. Unfortunately however, the level of organisational turbulence in primary care at the time of the study precluded their involvement. At the time of data collection (April 2002), many primary care groups (PCGs) were moving to primary care trust (PCT) status and effectively becoming new organisations. Reluctantly, it was felt that it would not be meaningful to include such brand new organisations in the study.

Despite a relatively high response rate of 61.4% from a sample of 1916 respondents, the perceptions of non-respondents remain unknown. Thus it is equally plausible to speculate that non-respondents held more or less favourable perceptions of progress than those of respondents. However, there were at least 10 respondents from each of the trusts in the sample, providing data on a range of perceptions from each study site. Most of the reported results are sample-wide, although reporting of the difference in domain scores between trusts required aggregation of individual respondents by trust. Such aggregation leads to dilution of differences of perception between different respondents within the same trust. Where perceptions of progress are widely divergent between respondents, the overall trust

score may be a poor indicator of perceptions. In order to assess the validity of aggregating scores and moving from individual- to trust-level as the unit of analysis, an assessment was made of the level of within and between trust variance using one-way analysis of variance. As scores varied more between trusts than they did within them, analysis at the level of trust is legitimate and results reported.

While each questionnaire was mailed directly to a named individual, there is anecdotal evidence that a number of individuals forwarded their copies to their trust clinical governance co-ordinators for completion. This became apparent during re-mailing of questionnaires to non-respondents, following contact from five individuals who had received more than one copy via colleagues within their trust. While it is not possible to identify exactly how many named individuals passed their questionnaire on to a colleague, it is likely that this is a relatively minor problem given the overall number of questionnaires and the limited number of times that that it was raised as an issue. One might argue that those study participants who passed their questionnaires to their trust clinical governance co-ordinator exhibit a lack of insight into their responsibilities for clinical governance.

Statistically significant results in the analyses of variance are tentative and should be treated with caution, as correlations between dependent variables gives rise to the potential for erroneous causal inference (Tabachnick and Fidell 2001). For example, if scores for quality improvement (DVB) and risk management scales (DVA) are highly correlated, and trust type (IVA) has an effect on quality improvement (DVB), it is likely that results of the effect of trust type (IVA) on risk management (DVA) will also reach significance. Yet, in

such circumstances, the apparent effect of trust type on risk management may be due solely to the variance shared by the two dependent variables. In contrast, non-significant results are robust i.e. where there is no effect between the independent variable and any of the dependent variables problems of confounding do not arise.

Notwithstanding the above caveats, results constitute a national overview of board and directorate managers' perceptions of current achievement against organisational competencies directly related to clinical governance in acute, ambulance and MH/LD/SC trusts. The following discussion draws out the implications of the major findings organised under four main headings: the primacy of the assurance agenda; structural bias; the potential for ritualistic reporting; and the impact of CHI visit on perceived progress.

1. The primacy of the assurance agenda

The results of the whole sample provide a clear and consistent picture of perceived importance, progress and shortfalls in clinical governance, both at the level of individual competencies and domain aggregates.

In an absolute sense, all individual competency items were considered to be important. Out of a possible maximum score of 6, the mean score across all competencies was 5.1, and no item had a mean of less than 4.5. While the difference between the perceived importance of items were small, the most highly rated items were existence of a formal committee (item 35: mean = 5.59), executive director with responsibility for clinical governance (item 32: mean = 5.54), existence of a no-blame culture (item 22: mean = 5.48; item 23: mean =

5.44), and risk management (item 17: mean = 5.44; item 14: mean = 5.42). These items principally relate to the quality assurance agenda, including high level structures and risk management. Items rated least important were related to the quality improvement agenda, including benchmarking (item 2: mean = 4.51) and use of research evidence in quality improvement (item 6: mean = 4.63; item 4: mean = 4.70). The same picture emerges at the aggregate level of domains. Out of a possible maximum score of 10, corporate accountability scored highest (8.85; 95% CI = 8.77 – 8.93), followed by managing risks (8.65; 95% CI = 8.59 – 8.71), performance improvement (8.15; 95% CI = 8.07 – 8.23), leadership and collaboration (8.04; 95% CI = 7.96 – 8.12) and finally improving quality (7.66; 95% CI = 7.58 – 7.74). Thus corporate accountability was considered marginally more important than other domains.

While variations in perceived importance were relatively small, variations in perceptions of achievement at both individual and domain aggregate level were much higher. The mean score of individual perceived achievements ranged from 3.14 (95% CI 2.98 – 3.30) to 5.70 (95% CI = 5.62 – 5.78) out of a possible 6. Results were similar to those for importance to the extent that items with the highest scores related to clinical governance committee structures (item 35: mean = 5.7; item 32: mean = 5.52), collation of complaints (item 20: mean = 5.25), local arrangements for committees (item 36: mean = 4.95) and discussing risk data (item 16: mean = 4.7). This again relates most strongly to the controls assurance agenda. Respondents perceived much lower rates of achievement against collaborative working across health communities (item 51: mean = 3.14; item 52: mean = 3.32), benchmarking (item 2: mean = 3.2) and use of research (item 6: mean = 3.38; 4: 3.52), user

involvement (item 49: mean = 3.52) or action for quality improvement (item 46: mean = 3.53). At the domain level the picture is clearer still- perceived achievement scores out of 10 were highest in corporate accountability (8.10; 95% CI 8.00 – 8.30), followed by managing risks (6.81; 95% CI 6.73 – 6.88), improving performance (6.29; 95% CI 6.19 – 6.39), leadership and collaboration (5.61; 95% CI 5.53 – 5.69) and quality improvement (5.40; 95% CI 5.32 – 5.48). Thus at both levels of aggregation respondents perceived much higher progress in corporate accountability and risk management than in either leadership and collaboration or quality improvement.

The fact that the domains perceived to be most important are also the ones where there is greatest perceived achievement is reassuring to the extent that respondents could be said to be targeting their efforts appropriately and achieving positive results. The perceived primacy of the corporate accountability agenda is understandable in the context of a strong performance management culture. Perhaps more worryingly however, the results could also be interpreted to support concerns about the difficulties of tackling the long-term quality improvement agenda of clinical governance (Freeman *et. al.* 2001; Roland *et. al.* 2001; Campbell *et. al.* 2002). Results suggest that five years on, less progress is being made in improvement-oriented aspects of clinical governance than in assurance-oriented aspects, as evinced by lower perceived achievement against items relating to joint working, user involvement, benchmarking and use of research evidence.

Given small differences between item and domain scores on perceived importance and larger differences on perceived achievement, results at both levels of aggregation suggest

that the areas with greatest shortfalls between importance and achievements relate to the quality improvement and leadership agenda. Aggregation of item scores into domains revealed the picture most clearly; the shortfalls between importance and achievement lie in leadership and collaboration (2.43; 95% CI 2.35 – 2.51), followed by quality improvement (2.26; 95% CI 2.18 – 2.34), performance improvement (1.86; 95% CI 1.78 – 1.94), managing risks (1.84; 95% CI 1.76 – 1.92) and corporate accountability (0.74; 95% CI 0.66 – 0.82).

Results suggest that the policy is being translated through implementation in acute, ambulance and MH/LD/SC trusts - the performance management context is resulting in an emphasis on the assurance agenda. This is evident in the way that assurance items were perceived as most important and most achieved against. It suggests that efforts have been directed to where it was anticipated that they would be most needed- the assurance agenda, and consequently gives weight to the assertion that support and development efforts should now be concentrated on the leadership and collaboration and quality improvement agendas. Concerns may also be raised about other unintended consequences of the performance management aspects of clinical governance: structural primacy and ritualistic reporting.

2. Structural primacy

Consistent with earlier evaluations reported in chapter 2 (Campbell *et. al.* 2001; Walshe *et. al.* 2000), results suggest perceptions of good progress against the structural agenda of clinical governance. Perceived achievements against structures were high, with a mean of 7.19 (95%CI 7.11 – 7.27) out of a possible 10. However, much less impressive were

perceived achievements against processes with a mean of 5.97 (95% CI 5.89 – 6.05), and outcomes with a mean of 5.64 (95% CI 5.56 – 5.72). Thus perceived achievements were lower against the arguably more important process and outcome related items- this is important as the absolute level of perceived achievement against the process and outcome indicate the effect of policy on day-to-day work practices.

There is a danger of concentrating effort on the structural mechanisms for clinical governance – committees, policies and resources - rather than the substance of the reforms and their intended outcomes in terms of the way that clinical teams work together to improve service provision. It may be suggested that this is simply a natural ‘lag’, and that one might expect the more difficult quality improvement agenda to be addressed in due course. A less optimistic assessment might be that the performance management context of clinical governance is reducing the emphasis on quality improvement, despite the perception that process and outcome changes are more important. The latter seems more likely as the corporate accountability domain was scored highest in perceived importance. The danger is that attention is focused on the systems rather than the effects that the systems are designed to achieve, and the board may increasingly rely on these systems to assure that all is well. The relatively high perceived importance and relatively low perceived achievement against process and outcome related items means that there was a relatively high perceived shortfall between importance and achievement for process and outcome aggregates, at 2.20 (95% CI 2.13 – 2.27) and 2.45 (95% CI 2.37 – 2.53) respectively.

3. Ritualistic exchanges

Multivariate analysis suggested that across domains, perceptions of progress were higher among board level managers than directorate level managers ($p < .001$). This is consistent with evidence from broader literature that perspectives on organisational achievement vary across levels of management within organisations (Keeton and Mengistu 1992; Singer *et al.* 2003). Given that board members perceptions of achievement are optimistic when compared to others, the reliance on the self-report of CEOs for estimates of progress in previous studies of clinical governance is likely to have produced overoptimistic assessments. Such differences in perception pose certain dangers in the context of quality assurance, and trusts may be experiencing difficulties associated with reliance on audit based systems of quality assurance in their implementation of clinical governance (Power 2000). Our results are consistent with clinical governance committees providing a 'theatrical' function, reassuring the board that all is well, while allowing 'business as usual' at lower levels within the organisation (Stronach 1999).

4. Impact of CHI visits on perceptions of progress

It must also give some cause for concern that our comparisons of managers' perceptions of progress in clinical governance between trusts that had been, were being, or had yet to be visited by CHI found few differences. There are a number of caveats to the use of study data to assess the impact of CHI visits on perceptions of progress in clinical governance. While data was used to assess the impact of CHI visits on managers' perceptions of progress, the study was not primarily designed to answer this question. The primary questions related to managers' perceptions of progress at a single point in time (cross-

sectional design). Had the study been primarily concerned with an assessment of the impact of CHI visits on clinical governance, then a process evaluation of the effects of a CHI visit within multiple case study sites would have been more appropriate. The use of cross-sectional data to assess the impact of CHI within the present study poses the difficulty of the non-random nature of the timing of CHI visits. Recently merged trusts were given a years grace before being visited, and trusts with known difficulties were visited earlier. While this might increase the significance of the findings, it more likely reduces the validity of the classification system on which the comparisons were made, given the cross-sectional nature of the study.

Further, there are likely to be difficulties associated with the relationship between the OPGC schedule and the CHI review itself. It is probably a self-fulfilling prophecy that managers in trusts close to a review are sensitised to answer on those issues that are most easily measured as part of a CHI review. As few statistically significant differences were found on any of the dependent variables between CHI visit-status groups, questions may be raised about the sensitivity of the OPGC i.e. its ability to detect between-group differences. However, its sensitivity is attested by its ability to detect statistically significant differences between both different types of trust and board members / directorate managers.

For all of the above reasons, results should be treated with caution. Caveats notwithstanding, results suggest little effect of CHI visits on most of the clinical governance competencies, especially those concerned with quality improvement and leadership & collaboration. The largely non-significant differences between CHI visit-

status groups on each of the three sets of dependent variables (domains; structure / process / outcome; and high / medium / low competency) is of particular importance given the robustness of non-significant results (Tabachnick and Fidell 2001). Specifically, results suggest a temporary and small rise in perceptions of achievement limited to performance improvement and corporate accountability domains. CHI has yet to be formally evaluated (Walshe 1999; Walshe 2002; Appleby and Coote 2002), and our findings suggest that process evaluation (Shaw 2000; Stecher & Davies 1987) is urgently needed to identify the conditions under which CHI reviews may lead to sustained quality improvement.

Conclusions

This chapter reported results of an application of the OPGC schedule in a stratified sample of 100 acute, ambulance and MH/LD/SC health care trusts in England. Substantive results showed that the competencies considered most important and recording highest attainment related to corporate accountability. The clinical governance literature clearly shows the juxtaposition of quality assurance and quality improvement at its heart; the results of this study suggest that to date, clinical governance implementation in acute, MH/LD/SC and ambulance trusts in England has favoured the former rather than the latter.

CHAPTER 9

CONCLUSIONS

“The recent explosion of performance targeting and monitoring suggests that policy authorities try to overcome (medical) resistance at the implementation stage by amplifying the scope and reach of performance management and creating new regulatory structures. Such responses, however, contain the seed of their rejection by clinical staff. ... clinicians of all disciplines are critical of what they perceive as a blind and unrelenting drive on increased throughput and reduced waiting times that inappropriately skew clinical priorities. Clinicians are also highly critical of the time now absorbed by ‘management generated and oriented (paper) work’ ‘totally lacking (in) clinical significance. These (often valid) conclusions provide grounds for further resistance.’”

(Degeling *et. al.* 2003: p.651)

Introduction

Chapters one and two identified the principal tension within clinical governance as the juxtaposition of internal continuous quality improvement and external quality assurance.

To better explore progress in clinical governance at corporate and staff levels, two new validated measures of progress in clinical governance were developed: the Clinical Governance Climate Questionnaire (CGCQ) and the Organisational Progress in Clinical Governance (OPCG) schedule. Chapters six and eight reported on the application of these measures, the former concerning staff perceptions of their routine work, the latter concerning perceptions of corporate progress. These results support a number of tentative conclusions concerning the effects of the improvement / assurance tension in clinical governance on implementation, the multiple nature of organisational climates and attendant difficulties surrounding attempts to influence them. A note of optimism is provided by the potential for dialectical resolution of professional differences relating to clinical governance.

OPCG results suggested that CHI reviews had a temporary and limited impact on clinical governance progress. However, as the study was not specifically designed to address this question, the impact of CHI reviews requires further evaluation. Five principal axes of evaluation may be discerned: intent; focus; perspectives; methodology; and data collection. Two 'ideal types' are then considered: the hypothetico-conductive paradigm comprising of summative, experimental and outcome elements; and the qualitative-naturalistic paradigm comprising of formative, naturalistic and process elements. Following a brief description, the major strengths and weaknesses of each are outlined in the context of large-scale health care programme evaluation. The possibility and desirability of hybrid evaluation designs is assessed, and the potential of formative, process oriented evaluation to identify the conditions under which interventions yield positive benefits and support improvement is highlighted. Such approaches may be particularly valuable in the context of the increasing integration of research provision with change and development under the Health and Social Care Research Governance Framework.

9.1 Progress in clinical governance: some conclusions

1. Staff perceptions of clinical governance climate(s) vary between organisations

While limited to four sites, the CGCQ study identified differences in staff perceptions of clinical governance climate between trusts. The PCT site consistently scored positively on clinical governance climate domains. This is consistent with the review of

implementation evaluations in chapter two, which indicated that PCTs have tended to approach clinical governance implementation in a facilitative way, harnessing clinical support. The variability between organisations suggests that a programme of organisational development could be used to improve clinical governance implementation through staff involvement.

2. There has been more achievement in quality assurance than improvement

The tension between quality assurance and quality improvement at the heart of clinical governance has resulted in implementation strategies favouring the former rather than the latter. OPCG domain aggregate scores indicate the primacy of the assurance agenda is in perceived *achievement*, with the highest perceived achievement in corporate accountability (8.1 95% CI 8.0 – 8.3) and the lowest in quality improvement (5.4 95% CI 5.3 – 5.5). As there was little difference between the perceived *importance* of domains, results suggest that the greatest shortfalls between importance and achievement lie in items associated with quality improvement, including joint working, user involvement and use of research to inform practice.

3. There has been more achievement in structural changes than process and outcome

OPCG results also suggest greatest perceived *achievement* in structural changes 7.19 95% CI 7.11 – 7.27) as opposed to processes (5.97 95% CI 5.89 – 6.05) or outcomes (5.64 95% CI 5.56 – 5.72), where outcomes refer to changes in the everyday work practices of staff. The concern is that the substance of the reforms may be being lost in the rush to make structural changes.

4. *Group and development oriented organisational climates are positively associated with clinical governance climate*

Construct validation results of the CGCQ against the Competing Values Questionnaire (CVQ) reported in chapter 5 were consistent with earlier studies, suggesting that group- and development-oriented cultures emphasising affiliation, teamwork, co-ordination and participation are associated with greater implementation of continuous quality improvement practices. In contrast, hierarchical cultures emphasising rules, regulations and reporting relationships are negatively associated with the implementation of quality improvement practices (Shortell *et. al.* 1995). This suggests that an overemphasis on structure in clinical governance implementation may be counter-productive to the long-term quality improvement agenda. It also suggests the need for organisational development in implementing clinical governance.

5. *Board level perceptions of progress are higher than those of other staff*

Multivariate analysis of OPCG results indicated that perceptions of progress were higher among board level managers than directorate level managers across domains ($p < .001$). The ability of audit based assurance systems to provide senior executives with the information needed to assess progress in clinical governance are thus called into question, and trust in such systems may be misplaced.

6. *Perceptions of clinical governance climate are mediated by professional affiliation*

CGCQ results reported in chapter 6 identified significant differences between professional groups for the domains of risk management ($F = 10.099$, d.f. = 3, $p = .000$), absence of unjust blame ($F = 12.479$, d.f. = 3, $p = .000$) and working with colleagues ($F = 5.110$, d.f. = 3, $p = .002$). On the former two, managers scored the clinical

governance climate more positively than doctors did. This suggests that perceptions of clinical governance climate on these domains are mediated through respondents' professional discipline.

7. Mechanistic organisational culture change is neither desirable nor possible

Attempts to develop a strong, unified organisational culture through transformational leadership are ethically dubious, given the tradition of moral concern inherent in clinical practice. Naïve cultural change initiatives imposed on clinical staff will meet active resistance.

8. Contingent, dialectical change may offer a way forward

Given the multiple nature of clinical governance climates identified above, there may be room for a dialectical resolution of conflicts between stakeholders, achieved by harnessing clinical and managerial perspectives in team-based approaches to clinical work.

9.2 Options for evaluating the effects of CHI reviews

OPCG results suggested that CHI reviews have a temporary and limited effect on progress in clinical governance. However, this finding is subject to a number of caveats related to the cross-sectional nature of the OPCG study - it was not designed specifically to assess the impact of CHI reviews. A more sophisticated option might be to design a process-based evaluation of the review process, drawing on multiple stakeholder perspectives and identifying the conditions under which the review process improves

clinical governance. The range of options confronting the would-be evaluator is wide and potentially bewildering:

"... there is no single, agreed upon definition for evaluation. Neither is there a single set of acceptable procedures that one follows in carrying out an evaluation study. Instead, there are a number of different conceptions about what evaluation means and how it should be done. We call these 'evaluation approaches', and we use the term to indicate a coherent set of ideas about what an evaluation should accomplish and how it should be carried out – in other words the goals and procedures of evaluation.."
(Stecher & Davies, 1987 pp. 22-23)

Pawson & Tilley (1997) list 'summative evaluation', 'formative evaluation', 'goal-free evaluation', 'theory-driven evaluation' and 'naturalistic-evaluation' among others.

Consequently, it is helpful to consider options under five axes (table 9.2.1): intent (summative Vs formative); focus (outcome Vs process); perspective (unitary Vs pluralistic); methodology (experimental Vs naturalistic); and data collection (quantitative Vs qualitative).

Table 9.2.1: Five principal axes of evaluation research

Intent	Summative (accountability)	Formative (improvement)
Focus	Outcome	Process
Perspective	Unitary	Pluralistic
Methodology	Experimental	Naturalistic
Data collection	Quantitative	Qualitative

Principal axes

1. Intent: summative (accountability) or formative (improvement)

Principally concerned with the question 'does the programme work', summative evaluations seek to make a judgment on the success of a programme (Clarke & Dawson, 1999). Typical examples of this approach include goal oriented (Stecher & Davies

1987), impact (Ovretveit 1998) and summative evaluation (Billings 2000). In contrast, formative evaluations explore how a programme may be improved. Evaluative approaches sharing this orientation include action research (Waterman *et. al.*) and user oriented evaluation (Stecher & Davies 1987). Summative approaches are particularly useful in the context of accountability to examine the extent to which a programme delivers its supposed benefits, while formative evaluations are better suited to educative applications, seeking to improve programmes over time. Both may make use of quantitative indicators; the former as 'dials' to assess impact and the latter as 'can openers' to figuratively lift the lid on underlying processes (Carter, Klein & Day, 1995). Summative approaches typically report findings at the end of the evaluation, delivering a final verdict. In contrast, formative approaches usually feed back findings regularly to participants throughout the life of the evaluation in order to facilitate improvements, particularly so in the case of action research. Guber & Lincoln (1989) provide a critical historical overview of the development of summative evaluation approaches via three distinct generations, against which they contrast their own 'fourth generation' approach.

2. Focus: outcome or process

While outcome evaluations assess the impact of programme interventions, process evaluations aim to understand the internal operation of programmes: *how* programmes produce an outcome under which specific circumstances, exploring key aspects of programme development in context. The former include goal oriented (Stecher & Davies 1987), impact and economic evaluations (Ovretveit 1998), and the latter theory driven (Chen 1990), realistic (Pawson & Tilley 1997) and implementation (Patton 1987) evaluation. As process approaches permit investigation of the informal, unintended

consequences of implementation, they are particularly useful for examining whether the programme is operating as intended, highlighting strengths and weaknesses and identifying areas for improvement. Predominantly drawing on qualitative case study designs, process approaches are concerned with identifying the reasons for programme success / failure in specific contexts.

3. Perspective: unitary or pluralistic

Unitary approaches, such as experimental and goal-oriented evaluation, consider programmes from the perspective of a single agency / professional group; all outcomes / indicators / objectives are identified by this group, typically that of the funding agency (Guber & Lincoln 1989). In contrast, pluralistic approaches such as user-oriented, responsive (Stecher & Davies 1987) and goal-free evaluation (Scriven 1972) share a concern to understand programme outcomes or process from multiple stakeholder perspectives. Emphases vary however; while responsive approaches are primarily concerned with understanding programme experiences from multiple perspectives (funding body, delivering agency, customer / patient), goal-free evaluations seek to move beyond 'official' goals to gather data on intended and unintended programme effects (King *et. al.* 1999).

4. Methodology: experimental or naturalistic

Experimental and quasi-experimental approaches have an empiricist focus, in which the social world is governed by general principles linked through causal chains. They apply controlled experimentation in order to derive generalisable conclusions about programme impact, either to assess whether outcomes are attributable to the

programme, or compare alternative approaches. This involves manipulation of the programme and its participants into controlled groups, in order to draw valid, replicable casual inferences. Examples include economic evaluations (Drummond *et. al.*) and experimental outcome studies. In contrast, naturalistic approaches study naturally occurring programme activities and process. Sensitive to departures from planning, unexpected variations and the vagaries of programme experience, they capture unintended programme effects as the programme unfolds. Examples drawing on naturalist designs include theory driven (Chen 1990), process (Billings 2000; Ovretveit 1998; Patton 1987), and decision (Stecher & Davies 1987) evaluation.

5. Data collection: quantitative or qualitative

Quantitative approaches involve collection of numerical data about aspects of the programme, via structured interviews, questionnaires / surveys or pre-coded observation schedules. In contrast, qualitative data collection involves recording people's experiences and the meanings that they attribute to events and behaviour via semi-structured interviews, focus groups, observation and / or documentary analysis (Murphy *et. al.* 1998; Runciman 2002; Pope *et. al.* 2002). Quantitative approaches are particularly suited to gathering data on a limited range of questions from large numbers of respondents. They do so by the application of standardised collection techniques, fitting responses into pre-determined categories. In contrast, qualitative methods are suited to gathering information from a small number of respondents in great detail, recording narratives rather than numbers. This facilitates exploration of competing explanations or understandings of programme operation and impact. Quantitative information is most suited to summative, goal-oriented, outcome and experimental

evaluation; qualitative to, responsive and goal-free evaluation. Other approaches, such as pluralistic, realistic and process evaluation, draw on both types of data.

Evaluation Paradigms

‘Textbook’ descriptions of evaluation outline two principal ‘pure’ approaches to evaluation: hypothetico-deductive and qualitative-naturalistic paradigms (Patton 1986). The former combines experimental design, quantitative data collection and statistical analysis; and the latter combines naturalistic methodology, qualitative data collection and content analysis (table 9.2.2).

(a) Hypothetico-deductive paradigm

Description

Derived from applications in the natural sciences, advocates suggest the suitability of experimental design to determine whether objectives have been reached, or to compare the effectiveness of alternatives to achieving a given objective. At its most basic, the approach requires randomisation of participants to intervention and control groups, specification of a hypothesis concerning programme effects, collection of quantitative data on specified outcome indicators pre and post intervention, and a statistical analysis that attempts to falsify the hypothesis. Possible designs include repeated-measures, matched-pairs and factorial. Where randomisation to treatment and comparison groups is not possible, *quasi-experimental* designs may be considered, incorporating repeated measures over time with a group, or over groups at the same time

Potential difficulties

Heterogeneity of intervention / context

The results of experimental programme evaluations often prove to be disappointingly inconclusive due to the complexity of the programme interventions themselves, and of the contexts into which they are implemented (Walshe & Freeman 2002). Interventions, especially in social programmes, are not standardised, making it difficult to make comparisons between programmes. Similarly, the contexts in which interventions take place will have a marked impact on the success or otherwise of the programme.

*Inability to identify *why* a programme works / fails*

In seeking to give an unambiguous answer to the question '*does the programme work?*' experimental designs neglect intervention processes, so are unable to answer *why* a programme succeeded or failed (Pawson & Tilley 1997). This is particularly important given the complexity of intervention and context outlined above.

Selecting satisfactory outcome criteria for assessing impact

There is a need for clarity and stakeholder agreement in selection of outcome criteria. This is often difficult to secure, especially when an evaluator wishes to include measures of unintended programme outcomes.

Table 9.2.2: Five principal axes of evaluation, by ‘ideal-type’ paradigm

Axis	Hypothetico-inductive paradigm		Qualitative-naturalistic paradigm
1. Intent	Summative “Does the programme work?” e.g. goal oriented, impact, and economic evaluation	Vs	Formative “How may the programme be improved over time?” E.g. Action research
2. Focus	Outcome “What impact does the programme have?” e.g. goal oriented, impact, and economic evaluation	Vs	Process “How does the programme produce results in specific contexts?” e.g. theory driven, realistic and implementation evaluation
3. Perspective	Unitary Evaluated from the funding agency’s perspective e.g. most experimental evaluations	Vs	Pluralistic Evaluated from multiple stakeholder perspectives e.g. user-oriented, goal-free and responsive evaluation
4. Methodology	Experimental “Are outcomes attributable to the programme?” e.g. outcome and economic evaluations	Vs	Naturalistic “What are the intended and unintended effects of the programme?” e.g. theory-driven, process and decision evaluation
5. Data	Quantitative Answers of many to pre-coded questions e.g. goal-oriented and outcome evaluation	Vs	Qualitative Meanings attributed by few respondents to events or behaviour e.g. goal-free and responsive evaluation

Establishing controlled conditions in social settings

Randomisation to control and experimental groups may prove difficult on ethical or practical grounds (Black 1996). The latter includes evaluation of existing programmes or where there is no comparison group. Under such circumstances, quasi-experimental designs may be useful.

(b) Qualitative-naturalistic paradigm

Description

Derived from anthropology and underpinned by diverse humanistic sociological perspectives, the unifying theme is that the study of humans is fundamentally different to that of the natural sciences; of primary importance are the meanings of human action and the situated contexts in which action occurs (Denzin 1978). Programmes are seen as dynamic entities, and the approach attempts to understand programme processes and their effects on participants in order to make improvements. In contrast to the hypothetico-deductive approach of the specification of hypotheses and variables prior to data collection, the endeavour is guided by a search for patterns. It is thus most useful in evaluations of innovative or emergent programmes, affecting a broad range of stakeholders. It typically involves a limited number of case studies, selected purposefully for their ability to shed light on a topic, using qualitative data collection methods in order to describe the case in detail and in context. Analysis is inductive, building theory and / or typologies from the data.

Potential difficulties

Lincoln & Guba (1985) identify a number of potential difficulties related to the reliability and validity of qualitative research methods in programme evaluation.

Internal validity (credibility)

This concerns the extent to which the subject was accurately described, and may be defended by use of prolonged involvement in the setting, use of a range of methods (triangulation), and attention to disconfirming cases in analysis.

External validity (transferability)

This concerns the extent to which findings may be generalised from the study to other contexts. As statistical inference is inappropriate outside of randomisation, recourse is made to rich description of the study case(s) in order to allow a reader to make a judgement on the transferability of results to other settings.

Reliability (dependability)

This is a necessary, though not sufficient, component of credibility. Based on the concept of auditing, it requires an examination of the methods used in the study. If they are considered in line with standards and acceptable, then the study may be considered reliable.

Objectivity (confirmability)

The substitution of 'confirmability' for 'objectivity' shifts the focus of enquiry from the researcher to the case study, and consists of an assessment of the 'fit' between the data and the findings, again through an assessment of the audit trail from data to results.

Beyond the two paradigms: eclecticism in design

Mixing different elements of methodology, data collection and analysis from the two paradigms has provoked considerable controversy (Murphy *et al.* 1998). Yet, many consider it possible and desirable to combine elements in order to satisfy stakeholder and decision-maker requirements. While exemplifying such an instrumentalist approach, Patton (1987) notes that the range of blends of methodology, data collection and analytical strategy is limited by certain constraints. These include the impossibility of converting quantitative data into detailed descriptions, or of creating an experimental setting and then studying natural evolution. Favoured combinations include the use of quantitative data and statistical analysis in naturalistic settings, and qualitative data and content analysis in experimental settings. For example, if study outcomes are best investigated by recording programme participants behaviour, experimental designs could and should use qualitative data collection and content analysis. Ultimately, good evaluation design depends upon an appropriate 'fit' between evaluation purpose, stakeholder requirements and available funds:

"The challenge is to find out which information is most needed and most useful in a given situation, and then to employ those methods best suited to producing the needed information." (Patton 1987:69)

Towards an evaluation of CHI reviews

While generally well received within the service, Day & Klein (2002) raise a number of important tensions within CHI clinical governance reviews. These include the summative and formative dimensions of the reviews; confusion over the balance between checking for the implementation of mechanisms or assessing their impact; queries over data quality and the extent of critical evaluation of data in context; all within the context of highly heterogeneous organisations. Review processes continue to evolve, albeit without the benefit of focused formative evaluation. Under the NHS Reform Bill, CHI's remit is set to expand from reviewing clinical governance arrangements to inspecting the service. It will become responsible for the 'star' system of trust assessment, and for overseeing a number of NHS inspectorates. Taken together, the new context and the experience of CHI's review methodology to date identify openings for a broad programme of evaluation. While by no means an exhaustive list, evaluations may provide some answers to the following questions:

1. Under what conditions do CHI reviews lead to practice change / improvement?
(process / realistic / theory driven)
2. To what extent have CHI reviews met the expectations of stakeholders; where do the difficulties lie; and how may they best be utilised in the future? (pluralistic / formative)
3. How do review teams actually go about their tasks? (Implementation)

Such a programme of research firmly places the emphasis on the difficulties and opportunities posed by implementation, and has the potential to act as a corrective to the external and summative assessment of progress based on the Performance Assessment Framework (PAF) and performance ratings ('stars').

Appendix 1: A digest of empirical studies of clinical governance implementation in primary and secondary care

This appendix provides an overview of the studies of implementation in clinical governance reviewed in chapter two. Methods and subjects are noted, and details of good progress, areas of concern and potential future difficulties are described.

Study	Method	Subjects	Good progress	Areas of concern	Potential future difficulties
Primary Care					
Campbell <i>et. al.</i> 2001a	Postal survey (Dec 2000)	Stratified (region) random sample of 72 (58, 81%) of 481 CG leads	. Shared learning . practice development plans 93% . PCG/T education 81% . Risk management 76% . Complaint analysis 50% . Significant event analysis 48% . Protocols / guidelines 45% . Benchmarking 41% . Use of incentives rising: 29% in yr1 to 50% in 2000		. Tension between engagement in QI and tackling poor performance, more acute with transfer of responsibility from Has to PCTs . Limited time and resources still a problem: 41% had no implementation budget
Campbell <i>et. al.</i> 2001b	Postal questionnaire to explore priorities, barriers and strategies (Dec 1999)	Stratified (region) random sample of 72 (52, 74%) of 481 CG leads	. 74% undertaken baseline assessment	. 45 (85%) had no HA support . Poor performance approaches largely educational	. Restrictions in resources and infrastructure likely to delay development
Campbell <i>et. al.</i> 2002	12 Qualitative case studies using interviews and documentary analysis: to explore role of CG	12 purposively sampled PCG/Ts	. Focus on developmental and supportive approaches. . Priority to establish sustainable infrastructures and implement 'bottom-up' quality improvement . Many of the approaches not new, but higher levels of corporacy . increased corporate identity of audit / risk management/ guideline activity . sharing of anonymised data for benchmarking purposes	. Barriers - pace of change - perceived blame culture - lack of staff - lack of funds - staff suspicion	. Challenge ahead lies in reconciling the dual role of supporting practices while monitoring and dealing with poor performance. . Excessive speed of change
Harrison & Lim 2000	Pilot semi-structured interviews	3 clinical governance leads			. Lack of hierarchical structures for enforcement of bureaucratic modes of

Study	Method	Subjects	Good progress	Areas of concern	Potential future difficulties
Primary Care					
					governance in PCG/Ts
Houghton <i>et. al.</i> 1999	Postal survey to explore the educational needs of PCG CG leads	Census of all 63 CG leads in the West Midlands region (57, 90%)		Tackling poor performance and leadership skills were the areas of practice where leads identified most need for training (mean 4+ on a 5-point scale).	
Marshall <i>et. al.</i> 2002	Semi-structured interviews and documentary analysis to explore the importance of culture and cultural change for CG implementation	50 clinical and non-clinical managers from 12 purposefully sampled PCG/Ts	. greater willingness of practices to work together / share information	. long history of GP independence . Political environment of delivery on targets . Lack of skills to make user involvement possible	Balancing specific tasks, such as implementing NSFs, with developing a climate of innovation
Preston and Baker 2000	Postal questionnaire to explore local strategies for CG and their needs	Census of CG lead in all 506 local PCGs in England and Wales (329, 65%)	. 178 (54.1%) had a developed a strategy, 138 (41.9%) were developing one . Strategies commonly featured audit and CPD activities. . Education activities to support CG reported by 267 (81.2%)	. 127 (56.9%) identified a need for training in change management; 133 (40.5%) in leadership skills; 130 (39.5%) facilitation skills . to develop CG culture, protected time identified as needed by 301 (91.5%); PDP by 274 (83.3%) and risk mgt by 265 (80.6%).	Support required in many areas of practice development (teamworking 200, 60.8%; information systems 248, 75.4%) as well as CG items
Roland <i>et. al.</i> 2001	Postal survey	Random sample (n=71) clinical governance leads			. Organisational context of I ⁰ not suited to hierarchical governance . Tension between improvement and assurance . Tension between non-blame culture and tackling poor performance . Short implementation timescale . Counsel for 'light touch' assurance and emphasis on cultural change
Spurgeon and De Luc 2002	Postal survey	Survey of CG 121 leads in PCG/Ts in West Midlands and South-west regions (70, 58%)	. All produced a documented baseline . educational programmes	. variability in time allocated for role, 23 (33%) with no formal allocated time	

Study	Method	Subjects	Good progress	Areas of concern	Potential future difficulties
Primary Care					
			used by 68 (97%) financial incentives used by 62 (88%)	barriers- - staff workload (65, 93%) - lack of financial resources - data quality (46, 66%)	
Taylor 2000	Semi-structured interviews	Systematic, non-probabilistic sample, matching 4 training practices with 4 non-training practices for size and geography	Teaching practices had a history of QI activity including audit and risk management in excess of CG requirements. These are a valuable resource for their PCGs	3 of the 4 non-teaching practices were willing to undertake QI activity but require support and leadership.	1 non-teaching trust was insular and threatened by CG- it was culturally ill-equipped to engage with CG
Wilkin <i>et. al</i> (2000)	Postal Survey	Stratified (region) random sample of 72 (53, 74%) of 481 CG leads		only 65% had identified a CG lead	
Wilkin <i>et. al</i> (2001)	Postal survey	Stratified (region) random sample of 71 (58, 81%) of 481 CG leads	100% had identified a CG lead. 93% are encouraging practice development plans 57 % have ½ day whole-practice teaching sessions 50 % were sharing information on care quality	Lack of time identified by 55% as a barrier 41% had no dedicated budget- the lack of capacity and resources poses difficulties	Collective learning shows evidence of some cultural change. But no evidence yet of improved care quality, and quality assurance (dealing with poor performance) not well developed) Increasing role of managing poor performance may lead to disengagement

Study	Method	Subjects	Good progress	Areas of concern	Potential future difficulties
Secondary, Community and Tertiary services					
Conduit <i>et. al.</i> 1999	Postal survey. Self assessment against 20 items taken from First Class Service	Census of 44 (40, xx%) CEOs of acute, community and ambulance trusts in Trent region.	Corporate accountability items (cttee structures and lines of accountability)	Quality improvement lower than other areas Clinical record keeping Clinical audit poorly	Balancing requirements for accountability with a 'no-blame' culture

Study	Method	Subjects	Good progress	Areas of concern	Potential future difficulties
Secondary, Community and Tertiary services					
Cowpe and Scrivens (2001)	Semi-structured interviews	40 medical / clinical directors from 5 acute, 3 MH and 2 community trusts	Greater willingness to address performance issues	targeted - changes in clinical systems and practices slow and uneven - Risk management not pursued with vigour in all	Lack of organisational capacity to manage the whole agenda
Franks 2001	Content analysis of trust CG reports to identify perceived priorities	All 21 annual CG reports from acute NHS trusts for 1999/ 2000 in Northern & Yorkshire region	Establishment of cttee structures, and lines of accountability	- User and patient involvement rarely mentioned - Difficult to discern use made of complaints info - Lack of good clinical data - Rarity of developing clinical leadership and team working - Poor integration of audit and risk management functions - Restructuring often superficial	- Rarity of organisational development approaches to tackle difficulties in a co-ordinated manner - Need to synthesise an organisational development approach to include users and carers
Freeman <i>et. al.</i> 1999	Postal survey. Self assessment against infrastructure, accountability and existing quality systems	Census to 39 (28, 72%) CG leads in acute, MH and community trusts in the SW region	- Formation of high level cttees - Lines of accountability	- Ambiguity over purpose and function of cttees - Amount of dedicated time for CG of CG leads - CG strategies developed by only 50% - Existing systems of QI / QA highly variable, requiring major work before meaningful integration	- Managers accountable for clinical quality failures but potential limited influence clinical practice - Collegiate nature of clinician influence sits uncomfortably with accountability - Balance between involving doctors and avoiding 'medical capture' - Cultural transformation
Freeman <i>et. al.</i> 2001a	Postal Survey. Self assessment of five existing systems of quality improvement against four dimensions of effectiveness	Census to CG leads in all 86 (74, 86%) trusts in two regions.	- Cttees to oversee audit / risk and complaints - Good coverage across directorates for complaints and adverse events. - Complaints and adverse incident systems considered effective	- Often no corporate structure for professional development and knowledge management - Patchy coverage in directorates for risk and knowledge management	Integration of systems requires thorough review and updating

Study	Method	Subjects	Good progress	Areas of concern	Potential future difficulties
Secondary, Community and Tertiary services					
Freeman <i>et. al.</i> 2001b	Semi-structured interviews examining intentions for measuring progress in CG	CEO, CG lead(s) and non-executive CG lead in all 47 trusts in the West Midland region (n=151)		17 (36%) trusts identified indicators relating solely to assurance, rather than the broader quality improvement agenda	Risk of jeopardising quality improvement through heavy handed performance measurement
Latham <i>et. al.</i> 1999	Postal survey. Self assessment against infrastructure, accountability and existing quality systems	Census to 47 (47, 100%) CG leads in acute, MH and community trusts in the SW region	- Formation of high level ctees - Lines of accountability	- Ambiguity over purpose and function of ctees - Amount of dedicated time for CG of CG leads - Existing systems of QI / QA highly variable, requiring major work before meaningful integration	- Managers accountable for clinical quality failures but potential limited influence clinical practice - Collegiate nature of clinician influence sits uncomfortably with accountability - Balance between involving doctors and avoiding 'medical capture'
Latham <i>et. al.</i> 2000	Describes combined results of Freeman <i>et. al.</i> 1999 and Latham <i>et. al.</i> 1999 above				
Walshe <i>et. al.</i> 2000	Semi-structured interviews	(a) Clinical governance lead(s), CEO, Non-executive director lead in all 47 West Midlands acute, community and MH/LD trusts (n=151)	- Committees and lines of accountability developed - Clinical governance leads are positive and committed to their role - Education and training initiatives used by many to implement CG	- While most reviewed existing QI structures, few made radical changes or integrated existing QI functions - Leadership contributions of CEOs limited to transactional rather than transformational styles; - No direct methods of indirectly or directly influencing staff behaviour-building a shared culture is being left to chance	- Barriers to CG include resource constraints- time; finance; poor information systems - while 'culture' discussed as something that may be manipulated, little evidence of change management plans with specific sub-groups 86 (48%) interviewees described measurement intentions solely in terms of assurance systems, raising concerns over perverse incentives
Wallace <i>et. al.</i> 2001a	Semi-structured interviews examining	CEO, CG lead(s) and non-executive CG lead in all 47 trusts in the West	Educational interventions common to raise profile of	Rare for CEOs to acknowledge a role for	Tension between learning from mistakes and acting

Study	Method	Subjects	Good progress	Areas of concern	Potential future difficulties
Secondary, Community and Tertiary services					
	vision for cultural change and the types of organisational interventions used and planned	Midland region (n=151)	CG (n=60)	themselves in leading organisational / cultural change (n=4). Much more common to describe systems and processes . No CG leads had concrete plans for leading organisational / cultural change . Limited number of respondents identifying participatory approaches such as guidelines (n=7) or peer review (n=9) . Some activity on performance appraisal systems (n=13) . Very limited application of benchmarking approaches (n=4)	decisively on poor practice . Lack of application of the principles of OD: articulating vision; building on existing values that support the new vision; aligning systems to reinforce changes; monitoring impact . Despite policy emphasis on culture change, no specific change models are identified
Wallace <i>et. al.</i> 2001b	Postal survey. Self assessment of use, perceived effectiveness and predicted future use of organisational strategies for influencing clinician's behaviour	Census to CG leads in all 86 (74, 86%) trusts in two regions.	. All trusts using at least 1 strategy . Educational programmes used by 96% of trusts	. Choice of methods related to previous local use, despite equivocal judgements of their perceived effectiveness	. Expectation of major cultural change through multiple optional means and then unspecified diffusion., rather than a systematic OD approach

Appendix 2: A digest of primary and secondary material on the use of performance indicators to increase health care quality

This appendix provides a digest of key findings from the review of performance indicators contained in chapter three. The appendix is subdivided into two tables. Table 2a focuses on general discussions of performance indicators, whereas table 2b focuses specifically on the use of outcome indicators as performance indicators. For each paper, the digest lists the study type, perceived aims of performance indicators, perceived problems, and perceived factors facilitating their use.

2a: Use of performance indicators in improving health care quality (general)

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
Audit Commission 2000 ¹	Discussion	1. improved public services 2. improved accountability 3. identify good practice and learn from others	Risk the process becomes an end in itself; manipulation of data increased if data published or reward and censure depend upon indicators.	1. clarity of purpose 2. focus on priorities 3. alignment with objective-setting 4. balance aspects of performance, costs of collection 5. regular refinement / continuity 6. robust and intelligible [attributable, responsive]
Audit Commission 2000 ²	Discussion			Robust indicators are: 1. relevant to organisation 2. clearly defined 3. easy to understand and use 4. comparable 5. verifiable 6. cost effective 7. unambiguous 8. attributable 9. responsive 10. avoid perverse incentives 11. allow innovation 12. statistically valid 13. timely
Barnetson and Cutright 2000	Content analysis	Increase objectivity and accountability by applying standardised indicators, and removing impartial bias	PIs are conceptual technologies that shape what issues are thought about by embedding normative assumptions into the selection and structure of indicators	
Barnsley J. 1996	Narrative review	Point the way to improved care processes; identify opportunities for clinical process improvement; identify further non-clinical opportunities to improve outcomes	Indicator validity (sensitivity and specificity); comparative validity (requiring adjustments for case-mix and hospital policy); Reliability (accurate, consistent coding) and definitional differences in eg 'surgery time' [open to close, or anaesthesia to exit of theatre]; Data availability and accessibility	Clarity on intended uses; commitment to common aim; involvement of clinicians in measures (to identify probs in their use); continuous feedback from users, incorporated into redesigned indicators; collaboration with records staff
Birch 1986	Discussion	Ensure effective operation [maximum return for minimum cost]; make explicit comparisons between the performance of providers or clinicians; highlight variation in resource use; provide objective support for existing suspicion	Need to control for exogenous and endogenous factors which affect the indicators [eg peculiar local supply factors; atypical demand factors]; specificity of indicators ['atypical indicator score is the result of poor performance'] is	

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
		and provide a trigger for action	improved by risk adjustment, as is sensitivity ['poor performance veiled by favourable exogenous factors']	
Blalock 1999	Discussion		Difficulty in attribution of effects to inputs; need to detail how implementation is affecting indicators; measures may not be highly correlated with aims and objectives; measures may lack validity ie not be the best proxies; there may be too few measures; they may become ends in themselves	
Boland and Fowler 2000	Review	Two axes of PIs are internal-external and formative-summative;	Feedback systems which punish poor performance may exacerbate poor performance	
Boyne 2000	Narrative discussion		PI regulation can only be justified if it delivers better services beyond what would have been provided in their absence	
Brown R.B. 1995	Narrative discussion		Perverse incentives leading to measure-fixation	
Carter N. 1989	Narrative review	Central control and accountability; provide unambiguous 'dial' of performance; act as 'tin openers' to a can of worms i.e. provide a partial picture for further discussion	Impact of environmental factors outside the control of the organisation; Conflicting objectives of health care require trade-offs; may be driven by current data rather than that required for specific goals	
Carter N. <i>et al.</i> 1992	Case studies	Prescriptive PIs- linked to targets (dials); descriptive PIs record changes (tin openers); proscriptive PIs specify what should not happen (alarm bells).	'Performance' is contestable and complex; use of prescriptive assumes clear standards are meaningful. Many used descriptively in practice as their meaning is disputed- requires persuasive management style	The acceptability of PIs to members- includes not measuring items dependent on external actors or environmental factors; accuracy and credibility of the indicators; bottom-up and peer review systems
Cave 1990	Case study	Facilitate 'management by exception' and allow 'hands-off' control [but this does not necessarily happen in practice]; 1. Accountability [external] 2. Control - to identify poor performance and lead to corrective action [chosen by users] 3. Development - part of a learning cycle [chosen by users]	Absence of a bottom-line means contentious issues between stakeholders are difficult to resolve; a 'natural cycle' of indicators in which need for accountability is flagged following a crisis. This triggers new indicators and a drive for improvement. To maintain improvement, managers have an incentive to manipulate ('game') the indicators. Attempts to correct this via proliferation of indicators leads to a complex system that falls into disuse, as attention shifts in the organisation; if PIs not set in appropriate context, they provoke hostility and are counterproductive	Solution to the problems of PIs are not just technical- PIs need to be integrated into working practices. If they are simply 'bolted on', they produce unintended consequences; clarity in the use of PIs [1 to 3 opposite]; need to specify and introduce system sensitively; i.e participation of users in formative stages; Conceptual framework for implementation is:- 1. analysis of environment 2. specification of indicators 3. integration with work practice 4. evaluation and change of PIs

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
Flynn N. 1986	Narrative discussion			Integrate measures with the management processes and culture of the organisation
Forster D.P. 1990	Secondary analysis of national data	Allows comparison of units / clinicians in a simple format	Measures may have low specificity (flag up potential problems that do not exist on closer examination)	
Goddard 1999	Interview	External accountability; internal management control		
Goddard <i>et. al.</i> 2000	Case studies		Feedback systems fail to recognise complex adaptive nature of systems; Unintended consequences echo Smith (1995); agencies rely on collaboration with other bodies to achieve outcomes	Involving staff; constant review; use of 'soft' data also to contextualise;
Goddard M. 2000	Qualitative interviews		Imperfect data will have dysfunctional consequences- measurement may not capture all aspects of outcome, esp long term effects, leading to tunnel vision and myopia; attribution can be difficult as some outcomes the result of co-operation, but incentives may not reward this, leading to sub-optimization, and reported behaviour may changed due to misrepresentation; reward structures may stimulate gaming by altering behavior for strategic advantage	Involve staff in the development and implementation of PIs; using other control mechanisms such as mutuality; measuring satisfaction; seeking expert help in interpretation; develop benchmarks independently of past activity
Goldstein H.			impossible to know whether a poor rating is due to something other than chance (specificity), or a hospital with a real problem has been missed (sensitivity)	

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
Greene J.C. 1999	Narrative discussion	Response to public loss of faith- PIs are visible and concrete; opportunity for collective decision on what outcomes to pursue; illuminate whether investments are adequate to achieve expected results	At best inadequate, at worst distorted; Cannot capture the human experience of participation in care; obscure the values and intents of the inquirer; serve single interests rather than encouraging open dialogue; fail to promote ownership and agency among stakeholders; accept the stated aims of programmes uncritically	Top leadership (resources); personal involvement of line managers; participation of stakeholders in developing useful measures; technical assistance; Clarity over use of information- data quality will fall if used to make budget decisions
Harley 1988	Discussion	Highlight variations; 'direction indicators' i.e. flag potential problems	Imprecision in data collection; PIs in formal performance review can adversely affect coalitions and decision-making; fixation on measures rather than on quality itself;	
Harley M. 1987	Narrative discussion	Identify hospitals at risk of performance failure before failure occurs		
Hepworth 1988	Narrative discussion		Public sector performance is more complex to measure as there is no 'financial bottom line'. Further difficulties are the over-riding importance of political objectives; the existence of multiple and conflicting objectives; short political time horizons	
Holloway J. 1999	Case study		Approaches are only as effective as the people who apply them	High status champion;
Jackson 2001	Case studies		Three generally applicable technical problems: interpretation may vary between orgs; confounding with issues beyond the orgs control; failure to take account of the full range of issues involved;	
Jacobs and Manzi 2000	Focus groups	Measurement for external accountability	Pursuit of narrow departmental interests; concentration only on key targets; concern not to be 'unique' (low risk taking); deception (manipulation of data)	
Jones G.T. 2000	Questionnaire	Central control and performance management; used to prompt questions		Secure acceptance of clinicians and managers; corporate vision, shared by all staff; PIs directly linked to organisations goals;
Kazandjian VA 1996	Case studies		Do differences in rates reflect quality of care or quality of data? Is there a correlation between care process and the indicator rates? Will severity adjustment change the relative position	Developed information system; a non-confrontational information web; enhanced investigative capacity to identify reasons for recorded sub-

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
			of organisation to its peers?	optimal performance; use of multi-disciplinary 'performance improvement teams' to discuss problems
Kouzman A. 1999	Case studies		Difficulty operationalizing an abstract concept, no simple 'bottom-line measure in the public sector;	
Leggat 1998	Systematic review	Improving operational effectiveness; external accountability [political, commercial, community]; fostering collaboration; benchmarking between organisations	1. content and quality of data; [Sensitivity and specificity of indicators, case-mix] 2. may be used to blame rather than to learn and correct- wrong as indicators are not proof of a problem, nor its solution	Structure indicators in relation to org strategy; indicators to provide information to monitor critical processes; use structure/ process / outcome indicators to enable connections between indicators and activities to be made; balance need for multiple measures with selection of only those critical to monitoring and adjusting; ensure data quality, involve stakeholders in development; deploy throughout the org.
Lowry 1988	Discussion	Identifying outliers for further examination	Incomplete and inaccurate data; time lags between data collection and publication of results; potential for misinterpretation;	
Marshall 2001	Narrative discussion	Ensuring accountability, stimulating quality improvement; external reporting may encourage development of internal data for QI processes		
McKee M. 1997	Editorial		Is it meaningful- manipulation of data ; Will it lead to improvement-Unintended consequences- less willing to operate on high risk cases,	
McKee M. 1998	Editorial		Distortion of behaviour in unintended ways, create pressures to be seen to do something without addressing real issues	
McKenzie F.C. 1998	Narrative discussion			Emphasize a few measures; focus on measures participants can control; balance accuracy and simplicity; align with culture of organisation; integrate with long-term incentives
McKevitt 1993	Discussion	1. institutional - focused on external stakeholders, concerned with renewal of legitimacy 2. managerial - focused on managers and professionals, concerned with internal		

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
		resource allocation 3. technical - focused on clients and concerned with evaluation of effectiveness / quality		
McLaughlin <i>et. al.</i> 2001	Discussion	Improving performance and health care quality	Difficulty of generating reliable and valid clinical data. Tension between meaningful data for national accountability and also useable to promote local quality improvement Use of Pis in the context of accountability leads to 'gaming'	
Meekings A. 1995	Narrative discussion	Fostering TQM: Prediction and insight; provide feedback; focus on learning not blame; alignment of objectives; involvement of staff.		Right implementation process: anticipating resistance; securing ownership; establish review process; develop top-to-bottom links; allowing for statistical variation; building balanced set of measures
Mulligan 2000	Discussion		Unable to distinguish between good and bad performance: unclear how aggregated indicators help to identify which parts of health systems contribute most to improved health; not clear to what extent indicators identify the contribution of health services to health- we do not know whether high death rates are due to poor care	
Nutley S. 1998	Review	External control / public accountability; identify poor performers and centres of excellence; to aid purchasers / patients to select providers; to enable providers to focus on improvement; public health data	Performance is multi-dimensional; different stakeholders have different information needs; many outputs are qualitative; PIs affected by measurement timing; PIs susceptible to manipulation; rely on goodwill of clinicians; 'backward looking' and a poor motivational device for action and change	
Nyhan R.C. 1995	Narrative discussion	Fostering TQM;	Inability to compare measures across disparate programmes	Developing objectives via consensus building; using objectives as an a priori basis for measurement; Employee involvement in selecting measures; feedback should be non-threatening and learning focused, not 'blame oriented'.
Page C.K. 1999	Narrative discussion	Fostering TQM within shared governance; Pis to aid in 'systems' thinking		Shared vision of all within the system; integration of strategy, design and

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
				outcomes with vision; engagement of employees and a 'team' approach;
Pollitt C. 1985	Qualitative interviews	Objective measurements needed for accountability of proper use of public funds; variations become accountable; legitimates acting on problems		
Popovich 1998	Discussion	Providing information for purchasers and consumers to judge quality of care		As performance is multi-faceted and may be viewed from professional, purchaser and consumer perspectives, each of these perspectives should be included
Roberts A. 1994	Narrative discussion	Monitor and control activities	Unintended side effects	Management control models must integrate planning, control and reward systems. Link performance measurement and reward systems to overall strategy
Schryve 1995	Discussion	Accreditation systems have 4 purposes; 1. discernment; results and processes both need checking to see if good outcomes mask poor processes, or poor outcomes result even from good processes 2. Prediction; of future performance is a function of current results and future management intentions 3. Improvement; fostered via guidance and measurement, and external comparisons 4. Recognition; of high standards via feedback		1. relevance; ability to measure important things 2. reliability; of collection 3. discriminatory power; ability to identify real variations 4. Risk adjustment; for confounding factors to facilitate comparisons 5. Validity; usefulness for purpose
Sheldon T. 1998	Narrative review	External verification of improvements in care quality	Displacement of informal modes of quality assurance; foster fear among professional staff; unintended consequences; data easily manipulated; case-mix variation; lack of adjustment data; Unintended adverse events: Short termism; measure fixation; gaming; manipulation; macro nature of indicators and micro nature of clinical decisions	Integrated and coordinated with other quality related systems eg guidelines; related to service frameworks; use local knowledge in interpretation; voluntary and non-judgemental; positive feedback of performance data;
Smith H. 1997	Narrative discussion	Assist policy formulation and service planning; improve service quality; review resource allocation; increase control and influence in decision-making	Minimum standards perceived as 'the norm'; Focus on failure may lead to fear of reporting	Standards based on users / clinicians expectations and experiences. Involve users in monitoring and review
Smith P. 1990	Narrative discussion	To overcome information asymmetries in the principal-agent relationship; external	Ambiguous targets create perverse incentives; long-term targets will not receive due attention;	

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
		accountability	responsiveness to local preferences means orgs may be pursuing different objectives; areas following same objectives may hve profoundly different needs; difficulty of achieving sensitive rewards / sanctions to give incentives for agents to respond	
Smith P. 1995(a)	Narrative review	Control of information asymmetry between principals and agents; alter processes via feedback of results into system	Problems due to mismatched principal requirements and agent responses and naïve production model; (a) divergence between objectives and measures- tunnel vision; sub-optimization; myopia; (b) inability to measure complexity- measure fixation; misrepresentation; (c) inability to process data correctly- misinterpretation; gaming; (d) ability to respond to new circumstances- ossification	Involve staff; retain flexibility in use of PIs; quantify every objective; keep reviewing system; measure satisfaction; seek expert interpretation; maintain careful data audit; nurture long-term career perspective; keep number of indicators small; develop independent benchmarks
Solberg L.I. 1997	Narrative discussion	To underpin QI; information for increased accountability	Accountability focus may result in defensiveness around case-mix and data quality	Clarity of use (improvement, accountability or research) ; not externally reporting 'improvement' related data; understanding differences between improvement accountability and research uses
Stewart 1994	Narrative discussion	Concern for management accountability and public accountability	Difficulties of measurement are inherent in the nature of public sector performance- there are no technical solutions. There is no financial 'bottom line'; complex multi-dimensional measures are needed, but value judgements still need to be made- performance is inherently contested. Legitimate political debates surround the definition of 'appropriate' measures- eg should social factors be taken into account? Relationships between inputs, outputs and outcomes are not necessarily clear.	
Tan 1999	Discussion			Achieve strong support from top management; involve staff in design of indicators; translate data into the orgs. management information infrastructure;
Tarr 1995	Discussion	Internal: 1. controlling and redirecting individuals [not easy when difficult to measure] 2. feedback to adjust performance or targets 3. performance against strategic goals. Can adjust performance or goals	Unintended consequences e.g. if the measure is indirect, the measure may become a substitute for the goal;	Develop PI systems away from command-control and use relative, statistical and other 'soft' data; design system as part of the corporate strategy; align systems and methods with corporate values; use 'vector' measures [direction and speed of change];

Paper	Study type	Perceived or reported general aims / advantages of PI systems	Perceived or reported general problems with PI systems	Perceived or reported factors facilitating the use of PI systems
				measurement seen as information not control; leave room for management judgement; constantly re-evaluate systems (values , goals and variables)
Thompson and Lally 2000	Narrative discussion		Validity, comparability; usefulness; distortion of activity; gaming, perverse incentives	
Van Peursem 1995	Systematic review	There is social value in disclosure of PIs	They are deceptive as they falsely convey an impression of 'objective truth'; by making some aspects visible, they marginalise other aspects and perspectives; Social outcomes can not be measured in precise terms, which creates problems for accountability (comparisons become difficult and potentially deceptive); Further measures are themselves only proxies for underlying concepts;	Offset precision and reliability issues by a holistic impression derived from a wide range of indicators; Reported measures should be seen only as indicators, focusing attention on issues of interest
Wellard 2000	Narrative discussion	Provide a starting point for discussion	May be used simplistically; measure-fixation, rather than the associated objective; tunnel vision, pursuing only those things measured; confounding of socio-economic characteristics or impact of other agencies leading to unwarranted praise or blame;	
Wyszewianski L. 1988	Narrative discussion	Facilitates quality assurance	Emphasis may be placed on measurement rather than action to improve	

Appendix 2b: Use of performance indicators in improving health care quality (outcome indicators)

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
Al-Assaf 1996	Narrative discussion	Outcomes are the result, so should be included alongside structure and process. They are also patient centred.		Collaborative effort of patients, providers, payers and regulators in forming guidelines; standardised outcome measures; efficient and error-free collection of data; automated data collection, input, analysis and retrieval / dissemination; decision support systems.
Ashton 1996	Meta- analysis		Need to show demonstrable link between indicator and the process of care [specifically prior stay for readmission]	Clear conceptual framework identifying confounding factors and taking these into account within the model
Blumberg M. 1986	Narrative Review		To avoid confounding, risk adjustment models are required if routine data is used. Risk adjustment models are themselves problematic.	
Brown R.B. 1995	Narrative discussion		Need to isolate confounding factors. Long-term therapies have many interventions so are complex. Measures confounded by the social and service circumstances or care	
Bull 1994	Retrospective cohort analysis		SMR for cervical cancer is a poor indicator of the quality of health services. Independent risk factors for mortality included smoking history and stage at diagnosis. Stage at diagnosis was associated with age and absence of smear history. District SMRs were not related to 5 year survival, stage at presentation of screening history- SMRs are thus subject to influence from a range of unidentified factors, and are too broad an indicator to be a valid measure of performance	
Collopy 1998	Discussion	Provide 'pointers' to the quality of care	Need to be educational- linking funding to results may have a negative effect on change through the 'gaming' of data	JCAHO considers following attributes of PI systems essential: process or outcome measures; current, continuous, automated database; data quality meets validity requirements; risk adjustment for confounding factors; timely feedback

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
Coyle 1999	Discussion			Knowledge of the connections between process and outcome is required to develop valid quality of care measures.- a conceptual framework is required. A framework for development incorporating Donabedian's structure-process-outcome model is discussed. Risk adjustment based on identified antecedents of care ensures results are 'medically meaningful'.
Davenport 1996	Cohort Study		Case-mix has a crucial effect on outcome data, particularly in non-randomised comparative studies	
Davies 1997	Review			
Davies H.T.O. 1995	Editorial	Intuitive appeal; barometer for quality of care	Interpretation- comparisons bedevilled by case-mix problems requiring large amounts of extra routine data; low validity and reliability of some measures; unblind assessment; insensitivity of outcome measures; do not provide information on what to change	
Davies H.T.O. 1998 (a)	Narrative review	Better targeting of resources; central to the assessment of quality;	Uncontrolled observations cause problems of interpretation due to confounding- need to adjust for case-mix; also contexts of care vary . Measurement problems are severe- possible options include QoI, well-being functional status etc. and choice of measure will determine rasnking; Blinding of measurement is not possible and will ensure measurement bias; Outcomes measures are relatively insensitive compared to process measures; Many inaccuracies in routine data; Weakness in producing action and change- unclear what action to take	Need to avoid confusion of evaluating therapies (research) with assessing quality
Davies H.T.O. 1998(b)	Narrative discussion	Quality improvements in health care	Meaning of measures- always remains uncertainty over performance and reported measures, leaving the possibility of false denigration or assurance; Effects of measures- may displace trust in professionalism to ritualistic 'systems'	

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
Davies H.T.O. 1998(c)	Narrative discussion	Control and accountability	End-of -process error detection focus which takes time to collect data and act on findings; methodological problems of weak ambiguous evidence; behaviour may be negatively influenced due to adversarial environment	
Donabedian A. 1992	Narrative discussion	Poor outcomes can identify cases that merit further analysis of process and structure in search of possible causes- they are a 'warning bell'. They also reflect skill in execution as well as appropriateness of choice of strategy	They conceal the detail needed to show precisely what went wrong- measures of structure and process are required for this. Open to misrepresentation and misunderstanding if multiple causation is not understood	Outcome can only be used to assess quality to the extent that structure, process and outcome are causally related. Recognition of the need for a range of structure, process and outcome measures
Giuffrida A. 1999(a)	Secondary analysis of routine data on admission rates		PIs should relate only to those factors under control of those it is applied to; Adjustments need to be made for population characteristics and supply of secondary care resources; year-to-year variation in rankings is high so measure unreliable;	Standardizing for socio-economic and supply conditions, as well as demographic factors
Giuffrida A. 1999(b)	Secondary analysis of FHSA admission rates		Impact of socio-economic factors on admission rates and rankings is considerable, as is secondary care supply;	
Giuffrida <i>et. al.</i> 2000	Secondary analysis of routine data	Measure aspects of performance for which organisations can be held to account;	Unintended consequences for behaviour; socio-economic confounding; influenced by the performance of other agencies in the health economy	
Greene J.C. 1999	Narrative discussion		Care quality as a manifestation of human experience is complex, an cannot be reduced to simple end points; standards for judgement of quality are pluralistic between stakeholders and cannot be captured in one measure; Judgements of quality is not meaningfully presented as 'objective', but must capture the dialogic potential of judge and that judged	
Gross <i>et. al.</i> 2000	Secondary analysis of routine datasets		Clinical indicators derived from different measurement systems gave different results of comparative performance. Confounding factors make comparisons between units difficult	
Hannan 1996	Secondary analysis of cardiac surgery	May be used to underpin QI initiatives, leading to improved outcomes (avoidance of mortality)	Belief that surgeons may refuse to operate on high risk patients assumes either	

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
	outcomes		1. Model predicts fewer deaths than occur in high risk patients 2. If surgeon experiences higher than average mortality in these patients. Hannan found that their model did not predict fewer deaths than occur in high-risk patients, and argues that to the extent patients are referred on to physicians with better outcomes records, the public is served	
Hendryx 1995	Secondary analysis		Small samples sizes of 'cases'; case mix; compare like physicians with like; Profile of both clinical and economic performance needed; which cases to compare?	Small samples sizes of 'cases' [collect over long time; aggregate to related diagnostic groups; study case types which are the majority of a physicians cases]; case mix [adjustment models]; compare like physicians with like [use physicians at comparable institutions]; Profile of both clinical and economic performance needed; which cases to compare? [high volume and high risk].
Hofer T.P. 1996	Secondary analysis of routine data, using a simulation model	Providing consumers with information on which to choose between rival providers	Poor specificity of mortality outlier data to correctly identify 'poor' hospitals leading to ¾ of hospital so identified actually being average	
Holland W.W. 1994	Narrative review	Avoidable mortality provides a warning signal ('sentinel event') of possible deficiencies- a starting point for detailed local inquiry	Mortality not appropriate where services focus on improving QoL or pain reduction.	Need adjustment for environment and social factors
Jenkins 1990	Narrative review	Monitor and evaluate progress towards more effective healthcare; evaluate efficacy of interventions; improve resource allocation; clinicians to evaluate their effectiveness and areas for improvement; monitoring of quality of provision		
Krivenko C.A. 1994	Secondary analysis and Qualitative interviews	Help understand the processes of care contributing to different clinical outcomes; identify best practice arrangements for wider dissemination		
Leng G.C.	Secondary analysis of routine data		Definitional problems of emergency (non-elective) readmission (second within a certain time point); large variations by specialty show	Outcome indicators should be direct objectives of care rather than proxies; severity of disease should be taken into

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
			importance of standardization for case mix	account
Leyland 1991	Secondary analysis of routine information		Crude 'rates' provide no indication of what, if any, action is appropriate to improve healthcare.	Statistical modelling of outcome probabilities based on national data, adjusted for case-mix variation. [problem- residuals that cannot be explained by the model are due to 'other factors'- these may include care quality, or any other variable not in the model] Also only looks at norms, so poor practice is included in the norming procedure...
Leyland A.H. 1997	Retrospective analysis of inpatient hospital records		Need to take account of case-mix, local social and environmental characteristics, and distinguish them from factors under the control of clinicians.	Multiple level modelling allows (a) Identification of specific aspects of performance (b) Quantifies unmeasured effects, so shows variation that cannot be explained by available data (c) Takes local circumstances into account; often not adequate patient level data for this
Linder 1991	Survey	Create external accountability		Outcome measurement alone does not create accountability- it requires informed and open discussion of difficult issues with professional networks. In the absence of openness, there is no change.
Logan 1991		Wider Q of whether an intervention improves health belongs to a different level of evaluation, and requires trial design rather than routine data; adverse occurrences in small numbers of cases require individual case analysis to explore circumstances, not aggregation into population rates	Link between services and outcomes is tenuous given dependence of outcome on host and environmental factors; Difficult to use for routine service evaluation: confounding for social class in neo-natal mortality; large data requirements for showing a 'real' difference in rates;	
Mannion and Goddard 2001	Case study		Impact on quality improvement is limited by perceived problems in data quality, time lag between collection and presentation, weak incentives, and belief in the importance of informal information	
Mant 1996	Narrative review		Wide range of confounding factors between care and outcome: Admission practices; Case-mix; Co-morbidity; case severity. Incomplete	

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
			data makes adjustment difficult, and data quality may be low; poor sensitivity to allow discrimination between the quality of care provided in different settings; value is for evaluating the effect of processes on outcomes in clinical trials for wider dissemination, not assessments of quality between units	
Mant D. 1995	Narrative review and data simulation		Methodological difficulties include case-mix, case-finding, severity, co-morbidity and chance; poor sensitivity of disease specific mortality compared to process measures	
McColl 1998(b)	Case studies	May be used to identify priority areas for further quality improvement work	Indicators not able to generate specific proposals for change; may be easily misinterpreted and potentially misleading; timeliness and quality of data may be suspect; low attributable of indicators to health care quality;	
McColl A. 1998	Narrative discussion	May promote evidence based interventions if there is evidence that the interventions on which indicators are based lead to improved outcomes		Attributable to health care; sensitive to change; based on valid and reliable information; reflect important clinical areas; precisely defined; include a number of dimensions of care; use evidence-based indicators associated with clear health gain [what of areas with little robust evidence?]
McColl <i>et. al.</i> 2000	Secondary analysis of routine datasets		Meaningful comparisons require : conformity of data collection; funding, which has opportunity costs	
McGlynn 1998	Narrative discussion	(US) Provide comparable data on care quality to influence consumer choice, purchaser decisions and internal quality improvement; provide a summary measure of the effectiveness of medical care; clearly represent the goals of the health care system	Tend to be influenced by many factors outside the control of clinicians; may be insensitive esp. if poor outcomes are rare and require large sample sizes; may take a long period of time to observe; may be difficult to interpret given length of time elapsed; modelling for interpretation	Include measures that satisfy multiple audiences; ensure process-outcome link has been empirically shown; ensure measures are under control of clinicians
McKee 1997(2)	Discussion		Is available data appropriate or of sufficient quality?; is it possible to adequately adjust for severity; are the numbers big enough to draw meaningful conclusions; [see McKee and Hunter 1995- answer is no to each]; perverse incentives created in both treatment decisions	

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
			and data recording; rankings not robust when adjustment data used [impossible to know how robust any ranking is]; manipulation of data; perverse incentives to select 'more viable' patients	
McKee M. 1995	Retrospective analysis of inpatient hospital records	Provide information to empower public and purchasers of health care	Disease category data is too small for meaningful conclusions; specialty rates need adjustment for case-mix including severity; coding quality for adjustment is suspect; interpretation involves detailed analysis and local knowledge of provision; publication of results may lead to actions not in the patients' interests; gaming behaviour such as episode inflation and other changes in recording and behaviour; publication of flawed data will increase confrontation and reduce co-operation	Need a severity adjustment scale based on UK data; FCE needs to be replaced ; increased record linkage
Mulley A.G. 1999	Narrative discussion	May be used to draw attention to issues that may need investigation	Many factors outside the control of clinicians affect measured outcomes; create incentives for surgeons not to operate on those who may benefit most	Evidence base for care elements associated with chosen measures; 'learning' approach rather than punishment and shaming; accepting that indicators are not robust enough to make judgements of care quality
Mulligan <i>et. al.</i> 2000	Narrative discussion		Confounding of socio-economic factors with episode outcomes. Sanctions thus penalise unfairly- if they are adjusted, they mask the extent of inequalities	
Nerenz D. 1998	Questionnaire		Weak statistical associations between process and outcome; sensitivity of measurement; costs of data collection; need for more sophisticated case-mix adjustment models; attribution difficulties, esp. in care of chronic illness	
Nutley S. 1998			Definitions of population ambiguous and may be affected by discharge policies; Seek to compare against a norm, but 'like for like' comparisons very difficult-as variances due to - clinical skill, available resources, case mix, data accuracy, chance variation; different adjustment models lead to different results	
Orchard C. 1994	Narrative review		Case-mix information is required on patient need and associated prognosis; prognosis is affected by non-medical factors; prognoses may	

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
			be difficult to attribute to specific treatments; effects of predictive factors vary by condition and between outcomes	
Parry G.J. 1998	Longitudinal study		Need to adjust for case-mix; uncertainty of estimates in small sample sizes; Rankings implied by league tables not justified and cause unnecessary blame and guilt; Annual league tables of mortality not reliable indicators as hospitals had wide and overlapping confidence intervals; any action prompted by annual league tables could equally be beneficial, detrimental or irrelevant	
Portelli 1997	Survey of 116 organisations in the ACHS Equip [Evaluation and Quality Improvement Programme]	Potentially stimulate a variety of quality activities- indicators accepted as 'flags' and lead to diagnosis and change [procedures; policy; technique; further quality activity]	Time needed to monitor and assess indicator data before issues responsible for elevated rates can be identified	Opportunity for clinicians to review indicators and discuss areas of difficulty; continuous monitoring and 6-monthly reporting to offer greater opportunity for problem identification
Proctor 1999	Discussion	External assurance; ethics- the need to balance use of resources with continuing advancement of care quality	Lower morale; shifting attention to the easily measurable	
Rigby <i>et. al.</i> 2001	Secondary analysis of routine data	Use of routine data to measure performance	Inaccuracy of routine data; confounding due to case mix variation between tertiary and district hospital admissions; confounding due to variation in patient selection	
Rissel C. 1998	Narrative discussion	Accountability of public funded services to taxpayers;	Lack of consumer input in developing indicators of outcome;	
Sheldon T. 1998	Narrative review		Chance variation effects; small numbers mean low significance; Outcomes affected by many variables not under control of clinician so difficult to attribute to performance	
Shiell 1997	Discussion	Reorient services so the spotlight shines on what is achieved rather than what is done	Outcomes are about values as well as technicalities- need agreed objectives between all stakeholder groups, and this is where the challenge lies	
Smith H. 1997	Narrative discussion		May be seen as an instrument of central control	Clinicians involved in defining goals and selecting indicators.

Paper	Study type	Perceived or reported advantages of outcome measures as PIs	Perceived or reported specific problems with outcome measures as PIs	Perceived or reported factors facilitating the use of outcome measures in PI systems
Smith P. 1995(b)	Case study	Increased control (accountability) over public sector organisations, both externally and internally	Rewards are dependent on data, which is in the hands of professionals so dysfunctional behaviour results; comparisons between units create incentive to be 'mid range' to escape scrutiny via behaviour, accounting or fraud; managers in inefficient units escape scrutiny because of favourable environmental circumstances; scope for interpretative misrepresentation is high	
Smith P. 1995(c)	Narrative discussion	Reflect the outcome for society of the public sector intervention; to inform political debate by providing stakeholders with information	Outcomes may be produced jointly with other organisations; Organisational environments differ; many outcomes only occur over long time scales; random variability in measures; reflect only some stakeholder interests; Data relied on is under control of staff they are used to control	Can only ever be used to 'hint' at outcomes emerging from activity;
Thomas 1991	Secondary analysis of routine data	Unplanned readmission may be viewed as an adverse outcome of care.	Severity and complexity is an important risk factor; clinical and other risk factors are different for different diagnostic groups- these confounding factors need to be controlled	Condition-specific, severity adjusted models for measuring readmission risk; identifying patients with low risk on the model but a readmission; peer review of records to establish whether quality problems occur at higher rates for such patients than controls

Appendix 3: Raw data from the focus groups: domain 1 (adverse events)

This appendix contains abridged raw data collected from the focus groups used to generate the sub-themes reported in chapter 5, and is included as an example of the data used to facilitate analysis, showing storage and retrieval methods. The example shown is one of nine domains analysed, taken from the domain 'adverse events' and reporting only those comments listed as reflecting 'ideal' behaviour in this domain.

Searching for Identifier: I1

#1 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: Mistakes should be accepted and 17
discussed openly. 18

#2 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: Information should be passed to the 20
relatives - it should be open. 21

#3 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: If I made a mistake, I would accept 23
the responsibility of that - to 24
discuss what has happened. 25

#4 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: Mistakes should be treated as a 27
learning process - we would learn by 28
our mistakes to ensure that things 29
are not repeated again. 30

#5 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: All necessary parties should be 32
involved - we should accept that a 33
mistake has been made and be able to 34
be open about it. 35

#6 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: All areas would be aware of what 37
has happened, why it has happened and 38
how we can go about learning. 39

#7 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: Things happen because of situations 41
and we have to work out how the 42
situation can be avoided in the 43
future. It might be due to overwork, 44
it might be due to stress, pressure, 45
etc. 46

#8 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: You can find yourself in a 48
situation where you are working on a 49
case and need an ITU bed but there is 50
no bed. Pt can't be sent for so you 51
wait and you wait and wait and then 52
you get under so much pressure and 53
the whole day you are under so much 54
pressure and then a mistake can 55
happen. 56

#9 of 112 GEHCHOL I1 +adverse events+

IDENTIFIER: I1

I1: You can learn from other people's 58
practices as well. If there is a 59
particular problem then it might be 60
worth looking at other areas and how 61
handle situations. 62

#10 of 112 GEHCHOL I1 +adverse events+

#14 of 112 GECHOL I1 +adverse events+

IDENTIFIER: I1

I1: Investigations should happen away 84
from the incident - need to involve 85
others outside of your services as 86
well. You get a different viewpoint, 87
a global picture and other issues 88
surrounding the incident. 89

#15 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: All parties involved or related 16
would get together to instigate an 17
investigation into events surrounding 18
the incident. 19

#16 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: There would be an examination of 21
what went wrong, how it went wrong to 22
find the true cause; whether an 23
individual or a system. 24

#17 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: The initial response needs to be 26
structured; there should be an 27
incident reporting mechanism. 28
Everyone needs to understand how this 29
reporting mechanism works. 30

#18 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: Some initial damage limitation 32
should take place immediately - if 33
there is an obvious loophole in 34
practice that has allowed a mistake 35
to happen you need to cure it there 36
and then so that it won't happen 37
while you are looking at other 38
causes. It would be unforgivable to 39
allow the same thing to happen again 40
while you spent weeks pontificating 41
over whether you can change the 42
syringe or whatever. 43

#19 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: The investigation should initially 45
be no-blame; not to look for 46
scapegoats. 47

#20 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: Openness should be encouraged. 49

#21 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: The investigation should be opened 51
up widely; you can get blinkered too 52
quickly looking for cause and effect 53
- almost knee-jerk reactions happen. 54
The actual reason for why something 55
happened may be organisational. It 56
may be outside the immediate area. 57

#22 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: You have to establish whether 59
current procedures have been followed 60
in the first place. If they have and 61
something has gone wrong then there 62
is a real issue - if not, then you 63
need to know why not in order to 64
identify a problem. 65

#23 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: There are habits of practice among 67
staff; things become acceptable - 68
"custom and practice" and sometimes 69
it is these things that can cause 70
problems. They are also difficult to 71
change. 72

#24 of 112 GEHTRAU I1 +adverse events+

IDENTIFIER: I1

I1: The search for solutions shouldn't 81
stop on the doorstep - we need to 82
look wider across the organisation 83
and beyond to see what best practice 84
is elsewhere - a lot of learning can 85
be done outside the organisation by 86
seeing how others deal with 87
situations. 88

#25 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: There would be a reporting process 14
to capture what had gone wrong. 15

#26 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: Need to find out all about what 17
went wrong: who was present, what was 18
prescribed, by whom, etc. 19

#27 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: I'd be opposed to simply 21
identifying individuals - it is the 22
process as a whole that should be 23
involved. 24

#28 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: There should be a non-judgemental 26
atmosphere to allow reporting; to be 27
encouraged to admit making mistakes 28
and being open enough to apologise. 29

#29 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: You need to know that you will be 31
supported, not scapegoated - because 32
if that happens things go 33
underground. 34

#30 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: Incidents have to be investigated 36
to find out why things happened - 37
were people trained properly? Was 38
there enough staff? Etc. 39

#31 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: Need to examine which underlying	41
processes have failed, not just blame	42
the individual at the centre.	43
[content]	44

#32 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: There should be protocols and	46
policies available and staff need to	47
be familiar with them.	48

#33 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: Feedback on incidents - to find out	50
what has happened in other areas -	51
identify common themes and share the	52
lessons widely across different teams	53
- to change practice to avoid things	54
happening again.	55

#34 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: A responsibility for collating and	57
developing action plans for the	58
future.	59

#35 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: Support for patients and staff	61
involved in incidents.	62

#36 of 112 GEHUROL I1 +adverse events+

IDENTIFIER: I1

I1: Circumstances examined, staffing	78
load at the time was assessed, policy	79
& procedures reviewed.	80

#37 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: Not blame an individual but look at	13
the process.	14

#38 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: Something happening at the end of	20
reporting the incident.	21

#39 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: An open investigation of what went	23
wrong. The reasons why - the	24
fundamentals of why things happened.	25

#40 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: An acknowledgement that people can	27
make mistakes.	28

#41 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: Building an action plan to address 30
the problems identified by the 31
investigation. 32

#42 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: You should be putting something in 34
place to ensure that it doesn't 35
happen again. 36

#43 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: Checking to see that whatever is 38
put in place is carried out - 39
measurable targets. 40

#44 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: Some immediate damage limitation - 42
a fast response straight away. Do 43
something in the interim period to 44
lessen the damage. 45

#45 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: Check that resources are in place - 47
staffing levels and resources needed 48
to do the job. In an ideal world 49
these would be in place. 50

#46 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: To lower protectionism and vested 52
interests, someone from outside 53
should be involved in the 54
investigation - third party is good 55
as long as they have appropriate 56
skills to deal with it. So clinical 57
knowledge is important, managerial 58
responsibilities and an impartial 59
view - to prevent a cover-up. 60

#47 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: People feeling that they will not 62
be unjustly blamed - allowing them to 63
be open and honest. 64

#48 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: Ability to say "no" to things that 70
you are not prepared for, in your 71
professional opinion. As a 72
professional, you need to know that 73
you have gone the right road. 74

#49 of 112 GHH1 I1 +adverse events+

IDENTIFIER: I1

I1: If there is an incident, it should 95
be used in learning opportunities 96
where no-one is worried - we discuss 97
what went wrong and use this as a 98
learning experience. 99

#50 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: Problem should be systematically	13
investigated to find out why things	14
went wrong.	15

#51 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: There needs to be a blameless	17
culture within the trust to enable	18
people to be honest.	19

#52 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: You would feel able to report the	21
incident without feeling that the	22
whole world will be let loose on you	23
- and you would lose your job.	24

#53 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: There would be policies and	26
procedures in place.	27

#54 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: There would be equality between	29
professionals; nurses would not be	30
singled out while doctors are	31
protected.	32

#55 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: The whole process and circumstances 34
should be examined; the activity at 35
the time, etc. instead of 36
highlighting the error, the total 37
situation should be examined. 38

#56 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: If you take the incident out of 40
context, you are just going to 41
scapegoat somebody - no change will 42
happen. 43

#57 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: People who are involved should 45
receive support through the process - 46
not made to feel guilty until proven 47
innocent. 48

#58 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: Action should be aimed at making 50
systems changes that avoid the same 51
conditions arising again in the 52
future. 53

#59 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: The experience needs to be positive 55
and used as a learning opportunity. 56
People should not be penalised for 57
making errors. 58

#60 of 112 GHH2 I1 +adverse events+

IDENTIFIER: I1

I1: Procedures are not sympathetic to	60
the staff involved - there is no	61
support or looking for something	62
positive.	63

#61 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: We would review procedures and	13
practice to reduce the chances of the	14
error occurring again.	15

#62 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: We would be open about the event	17
and have a no-fault reporting system.	18

#63 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: We would encourage people to record	21
and report "near misses".	22

#64 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: The reporting system would be	24
"no-fault" based and learning from	25
incidents would be simply a normal	26
part of what we do.	27

#65 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Would check that any protocols that 29
were in place had been followed. 30

#66 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Generating new protocols in light 32
of the findings of an investigation. 33

#67 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: We have to be sure that we are not 35
just reporting; that we are also 36
acting on our reports. 37

#68 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Some changes have cost implications 39
- if it is a resource question, then 40
it becomes much more difficult to 41
make changes. 42

#69 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Senior managers would acknowledge 44
and accept the difficulties and risks 45
and provide the resources to 46
minimise them. 47

#70 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Was the error done by someone 49
operating at a level they were 50
competent to do. 51

#71 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Mutual support and willingness to 53
help. 54

#72 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: People would be open and say "well, 56
I did this: I wish I hadn't but I 57
did". 58

#73 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Disciplinary action may well result 60
from an incident - when that does 61
happen people see that it isn't a 62
no-fault system. 63

#74 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Information on the incident would 65
be collected first - different people 66
have a different idea of what the 67
facts are. 68

#75 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Someone with a good understanding	70
of the mechanics of the situation is	71
required to run the investigation.	72

#76 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: Investigations should be seen to be	74
independent and competent to	75
investigate.	76

#77 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: All of the associated circumstances	78
would be investigated and taken into	79
consideration.	80

#78 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: People will still feel guilty and	82
feel the need to talk things through	83
- there is a distinction between	84
responsibility and blame. People	85
should take responsibility for their	86
own actions - but there might be all	87
sorts of other things going on as	88
well.	89

#79 of 112 MSCONS I1 +adverse events+

IDENTIFIER: I1

I1: We should avoid thinking	91
automatically that there is someone	92
to blame.	93

#80 of 112 MCONSC I1 +adverse events+

IDENTIFIER: I1

I1: There is a system of reporting in 95
place that people know about. 96

#81 of 112 MCONSC I1 +adverse events+

IDENTIFIER: I1

I1: The reporting of adverse incidents 98
is encouraged. 99

#82 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: It would be reported, acted on and t 15
loop closed. 16

#83 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: It would be shared across the trust 18
even with other trusts, to ensure that 19
didn't happen again. 20

#84 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: Reporting would be via an incident 22
reporting system, robust enough to go 23
a clinical risk management system, or 24
trust directorate system. 25

#85 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: In an ideal world incidents would be 27
flagged up. 28

#86 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: Clinical Supervisor would pick up on 30
adverse incidents. We would discuss w 31
happened, why it happened and the team 32
would give support: emotional support 33
getting people to act as a team - if w 34
are a team, we should be there to help 35
another. 36

#87 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: There should be "no blame" - things 38
be looked at in the cold light of day: 39
there shouldn't be pointing of the fin 40
to start with. 41

#88 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: Things may have happened that requir 43
learning by an individual or others - 44
needs to be identified and provided. 45

#89 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: There has to be a difference with 47
negligence - if something is blatantly 48
negligent as opposed to an accident. 49
Accidents should result in a lot of 50
support to individuals - blatant 51
negligence would not lead to supportin 52
the person that has committed the act. 53

#90 of 112 MSSURG I1 +adverse events+

IDENTIFIER: I1

I1: There is a fine line between negligence
and accidents - human error and process
failure. 55
56
57

#91 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: Incidents are not necessarily a 13
point in time - they perhaps develop 14
over time. In an ideal world, we 15
would be able to spot them coming and 16
act in a way to lessen the impact 17
of them. If they are out of the 18
blue, need to deal with it in a way 19
which is least restrictive to all 20
concerned and supportive. While 21
thinking afterwards - what are the 22
things we could have done - so 23
reflection is important. 24

#92 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

I1: Supportiveness would be people 26
working together to try and solve the 27
problem rather than people going 28
'Oh my god, this is a problem' and 29
quickly distancing themselves in the 30
opposite direction, which is what 31
normally happens. 32

#93 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

I1: Also people being able to feel that 34
they will get practical support - 35
things done, from colleagues with 36
experience, but also emotional 37
support- people get stressed out, 38
worried and frightened, and that's 39
clients and staff. 40

#98 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

I1: There needs to be some ongoing	74
evaluation of how the first step has	75
gone - then keep reviewing progress	76
making new plans, and implementing	77
them. An on-going cycle really. So	78
evaluating them is important.	79

#99 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

I1: Ideally there would be people with	81
skills - approving how things had	82
gone - so some training implications	83
here, and resources.	84

#100 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

I1: We would need to identify	86
implications and act; either by	87
informal processes, or more formal	88
processes of reviewing incidents:-	89

#101 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

I1: Once the immediate crisis is over,	91
it is useful to sit down and review.	92
Ideally it should happen. This could	93
happen in a group, to get feedback	94
or how things went and what needs	95
to happen, in the context of our	96
team meetings.	97

#102 of 112 PREMMH I1 +adverse events+

IDENTIFIER: I1

I1: There needs to be openness - if 99
this is here now, it is ---- that 100
when things occur people will learn 101
from it - to anticipate things 102
happening in the future, before they 103
actually do. If we foster the 104
openness, we can learn from things - 105
training or resource implications. 106

#103 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

#-I1

I1: Team would be very supportive - 12
allowed to discuss the incident at 13
length, and discuss why things went 14
wrong. 15

#104 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: You should feel safe, and able to 17
access someone to discuss issues and 18
talk it through. 19

#105 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: Line management support as well as 21
colleagues. 22

#106 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: It would be documented as a 24
clinical incident - recorded. 25

#107 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: Some kind of analysis (if	27
possible) to avoid it happening	28
again. A lot of things that shouldn't	29
happen as a result of poor	30
communication.	31

#108 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: Need to relate 'problems' to the	33
wider context in which case takes	34
place without blame of individuals -	35
look at system failure rather than	36
blame the person.	37

#109 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: In the aircraft industry pilots	39
have a 'mistake/near miss of the	40
month' - they publish them and	41
discuss them at professional meetings	42
- I think we should have this as well	43
and it should be encouraged.	44

#110 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: You should be patted on the back	46
for coming forward with something	47
rather than trying to hushing it up.	48

#111 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: Needs to have a tight timescale -	51
discussed, enabled and finished quite	52
quickly and in good time.	53

#112 of 112 PREMN I1 +adverse events+

IDENTIFIER: I1

I1: Underlying causes being looked for	55
(systems) and changes made to the	56
underlying causes rather than	57
individual blame.	58

Appendix 4: The CGCQ validation study questionnaire

This appendix contains a copy of the CGCQ validation study questionnaire reported in chapter 5, section 5.1.5. The questionnaire was sent to all operational clinical governance managers at hospital, community, mental health and ambulance trust in the West Midlands region, between April and May 2001.

The Clinical Governance Climate Questionnaire (CGCQ)
A measure of progress in clinical Governance

Instructions to participants

The CGCQ currently consists of 29 areas of interest, organised under nine general headings. For example, the heading 'adverse events' contains four areas of interest, including 'openness and encouragement to report' and 'blame'.

Each area of interest has a number of statements (items) listed under the definition. We would like you to indicate how good an indicator of the area of interest each item is. For each of the items, say whether you agree or disagree **that the item is a good indicator of the area of interest**. Please rate all the items on a 5-point scale by placing a tick [✓] in the box best representing your opinion. Please answer all of the questions.

If you wish to suggest changes to the wording of any item, or suggest alternative items that you think better describe the area of interest, then do so in the box on each page.

On completion of the schedule, please return it by **May 3rd** to:

Mr. T. Freeman
Research Fellow
HSMC
Park House
40 Edgbaston Park Road
Birmingham B15 2RT

Thank you for your co-operation, it is much appreciated. Please do not hesitate to contact me if you have any questions or wish to discuss this further.

Tel:

email:

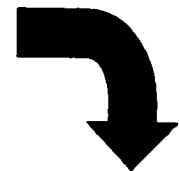
© Tim Freeman
HSMC 2001

For each of the items, say whether you agree or disagree **that the item is a good indicator of the area of interest**. Do not answer whether you agree or disagree with the item itself.

Please rate all the items on a 5-point scale by placing a tick [✓] in the box which best represents your opinion. Please answer all of the questions.

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



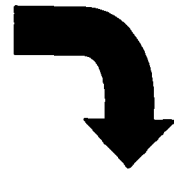
ADVERSE EVENTS: Openness and encouragement to report adverse events

	1	2	3	4	5
1. We are encouraged to admit our mistakes	☐	☐	☐	☐	☐
2. If I made a mistake I'd report it	☐	☐	☐	☐	☐
3. People here feel able to own-up to mistakes	☐	☐	☐	☐	☐
4. If someone makes a mistake, they try to cover it up	☐	☐	☐	☐	☐
5. If I made a mistake I'd be open and honest about it	☐	☐	☐	☐	☐
6. Only a fool would admit to making an error	☐	☐	☐	☐	☐
7. People are congratulated for coming forward about errors or mistakes	☐	☐	☐	☐	☐
8. People who identify adverse events are seen as just creating problems	☐	☐	☐	☐	☐
9. Error reporting systems are basically a stick to beat clinicians with	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



ADVERSE EVENTS: Blame

	1	2	3	4	5
1. When something goes wrong, there is a search for someone to blame	☐	☐	☐	☐	☐
2. If I made a mistake I would be supported	☐	☐	☐	☐	☐
3. There is no support for staff involved in errors	☐	☐	☐	☐	☐
4. When things go wrong, there is an automatic assumption that 'someone is to blame'	☐	☐	☐	☐	☐
5. If I made a mistake, I'd be unjustly blamed for it	☐	☐	☐	☐	☐
6. People involved in clinical incidents are made to feel guilty	☐	☐	☐	☐	☐
7. When there is an error, we look for failures in systems rather than blame individuals	☐	☐	☐	☐	☐
8. People involved in clinical incidents are made to feel like criminals	☐	☐	☐	☐	☐
9. The emphasis is always on how an incident it happened, not who made the mistake	☐	☐	☐	☐	☐
10. We work in an atmosphere of blame	☐	☐	☐	☐	☐
11. If something goes wrong, someone will be scapegoated	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

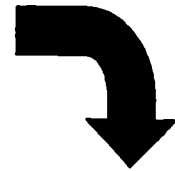
1= Strongly agree

2= Agree that the item is a good indicator of the area of interest

3= Neither

4= Disagree that the item is a good indicator of the area of interest

5= Disagree strongly that the item is a good indicator of the area of interest

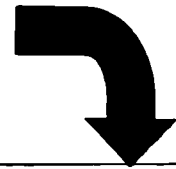


ADVERSE EVENTS: using a broad approach in investigations

	1	2	3	4	5
1. The underlying causes of clinical incidents are simply ignored	☐	☐	☐	☐	☐
2. When things go wrong, there is a thorough investigation to find the real causes	☐	☐	☐	☐	☐
3. After an incident, there is a wide-ranging investigation	☐	☐	☐	☐	☐
4. All of the circumstances surrounding a clinical incident are investigated	☐	☐	☐	☐	☐
5. When things go wrong, we discuss the incident at length	☐	☐	☐	☐	☐
6. We look beyond the immediate situation to find the real causes of clinical errors	☐	☐	☐	☐	☐
7. There is pressure to 'solve' problems quickly rather than take time and do it properly	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

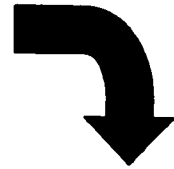
1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



	1	2	3	4	5
1. When things go wrong, we change practice to avoid it happening again					
2. After an incident, we change our systems to stop the same conditions arising again					
3. We've yet to change practice as a result of an investigation into an incident					
4. Risk assessment processes are updated in the light of incidents					

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



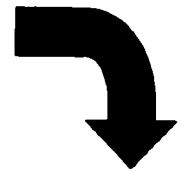
QUALITY IMPROVEMENT: Focus on organisational processes

	1	2	3	4	5
1. We always check the effects on others of any quality improvements we make	☐	☐	☐	☐	☐
2. When making improvements, we try to get at the root causes of problems	☐	☐	☐	☐	☐
3. In solving one quality problem, we create others elsewhere	☐	☐	☐	☐	☐
4. The first we know of quality improvements elsewhere is when we feel their effects	☐	☐	☐	☐	☐
5. Often, improvements made elsewhere make our work more difficult	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



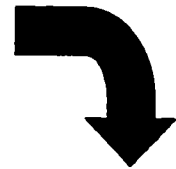
QUALITY IMPROVEMENT: Explicit focus on internal and external customers

	1	2	3	4	5
1. We involve patients in our quality improvement activity					
2. We consult widely with colleagues when deciding priorities for quality improvement					
3. Before making service improvements, we consult with patient representatives					
4. We agree our quality improvement activity with those affected by changes					
5. When improving services, we focus on service needs rather than patient needs					

Please make suggestions for re-wording of items, or offer alternative items, in the box below

[illegible]

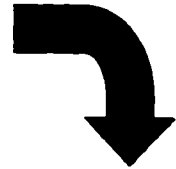
1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



	1	2	3	4	5
1. We work across teams to make improvements in whole systems of care					
2. There is no cross-departmental work to improve quality					
3. We often work together across teams to improve our services					
4. People seem to fight their own corner rather than work together					
5. We tend to improve our own areas, rather than work together to make big changes					

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



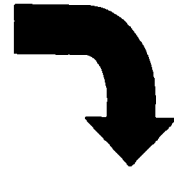
QUALITY IMPROVEMENT: Investment in improvement activity

	1	2	3	4	5
1. Quality improvement is seen as an important activity	☐	☐	☐	☐	☐
2. We have protected time for quality improvement activity	☐	☐	☐	☐	☐
3. There is no time to reflect on practice	☐	☐	☐	☐	☐
4. We continually try to improve the quality of our practice	☐	☐	☐	☐	☐
5. Clinicians are made to feel guilty about time away from direct patient contact	☐	☐	☐	☐	☐
6. Immediate pressures are always more important than quality improvement	☐	☐	☐	☐	☐
7. Long-term planning for quality improvement gets lost in the day-to-day	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly

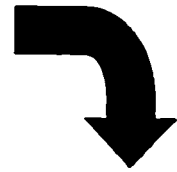


QUALITY IMPROVEMENT: Environment for improvement

	1	2	3	4	5
1. Quality improvement is an important part of being a professional	☐	☐	☐	☐	☐
2. Quality improvement is imposed from above rather than built from below	☐	☐	☐	☐	☐
3. I feel involved in quality improvement- we are changing rather than being changed	☐	☐	☐	☐	☐
4. People are encouraged to lead on improvement in their areas of interest	☐	☐	☐	☐	☐
5. I am motivated to improve quality	☐	☐	☐	☐	☐
6. If I had an idea for improving quality, I'd know how to take it further	☐	☐	☐	☐	☐
7. Our quality improvement activity is largely a response to external pressure	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



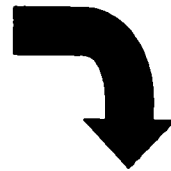
1. We use specific quality improvement techniques
2. We have no formal quality improvement processes
3. I've been involved in developing care pathways
4. When making quality improvements, we draw on a wide range of techniques

1 2 3 4 5

[illegible]

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly

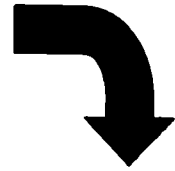


QUALITY IMPROVEMENT: Planning quality improvement

	1	2	3	4	5
1. In our trust, quality improvement is about the latest fad	☐	☐	☐	☐	☐
2. Around here, service improvements tend to be crisis-led	☐	☐	☐	☐	☐
3. We have a clear system for selecting which innovations we will adopt	☐	☐	☐	☐	☐
4. We base our improvement activity on an assessment of current problems	☐	☐	☐	☐	☐
5. We use research evidence to inform our improvement activity	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly

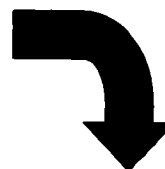


1. Clinical risk policies are shared throughout the organisation
2. Each area handles clinical risk in its own way
3. There is no common approach to risk management here
4. We try to spot potential clinical incidents before they happen
5. We react to problems, rather than try to prevent them
6. We don't look at clinical risks in a systematic way

1	2	3	4	5

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



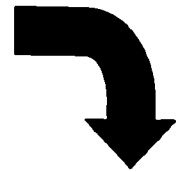
RISK MANAGEMENT: Assessment of risk

	1	2	3	4	5
1. We systematically assess clinical risks	☐	☐	☐	☐	☐
2. I'm conscious of the clinical risks involved in my work	☐	☐	☐	☐	☐
3. We collect information on clinical risks	☐	☐	☐	☐	☐
4. We have no formal risk assessment systems	☐	☐	☐	☐	☐
5. We collect information on the clinical risks that matter most	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



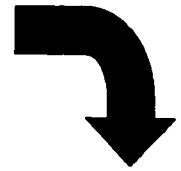
RISK MANAGEMENT: Action on identified risks

	1	2	3	4	5
1. When a clinical risk is identified, there is always action to address it	☐	☐	☐	☐	☐
2. Clinical risks are taken very seriously here	☐	☐	☐	☐	☐
3. Identified clinical risks simply remain undressed	☐	☐	☐	☐	☐
4. We react to incidents rather than plan to avoid them	☐	☐	☐	☐	☐
6. We don't address the accidents waiting to happen	☐	☐	☐	☐	☐
7. There is no sense of urgency in managing clinical risks	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly

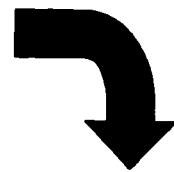


ORGANISATIONAL LEARNING: Investment in learning

	1	2	3	4	5
1. Time spent reflecting on practice is seen as 'not real work'	☐	☐	☐	☐	☐
2. There is no time to get together to share ideas	☐	☐	☐	☐	☐
3. We are encouraged to spend time sharing ideas with others	☐	☐	☐	☐	☐
4. We are encouraged to devote time to disseminating our good practice	☐	☐	☐	☐	☐
5. Reflection on practice is accepted as a valuable part of being a professional	☐	☐	☐	☐	☐
6. Sharing quality improvements is seen as an essential part of good practice	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

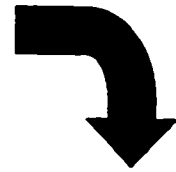
1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



	1	2	3	4	5
1. I don't know about good practice taking place in other parts of the organisation					
2. We share practice issues with others in different parts of the organisation					
3. People tend to hold on to information for their own advantage					
4. In this organisation, good practice stays isolated in pockets					
5. When we learn something useful, we share it across the whole trust					
6. Around here, teams from different parts of the organisation share their good practice					
7. We share good practice ideas with others outside the organisation					
8. There is support for sharing insights on practice across the organisation					
9. We are keen to learn from others outside our organisation					
10. We have a 'not invented here' attitude to new ideas					

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly

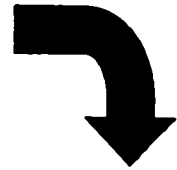


ORGANISATIONAL CHANGE: Motivation to change

	1	2	3	4	5
1. I can't see the benefit of making changes to services	☐	☐	☐	☐	☐
2. People here are free to question 'custom and practice'	☐	☐	☐	☐	☐
3. Most people here want to work in the same way that they always have done	☐	☐	☐	☐	☐
4. People here are highly motivated to make changes to clinical practice	☐	☐	☐	☐	☐
5. People here are prepared to try new ways of doing things	☐	☐	☐	☐	☐
6. There is a great deal of inertia which stops most changes to practice	☐	☐	☐	☐	☐
7. It is not my place to suggest changes to services	☐	☐	☐	☐	☐
8. We tend to do things as they have always been done	☐	☐	☐	☐	☐
9. We are forced into making service changes, rather than encouraged to make them	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly

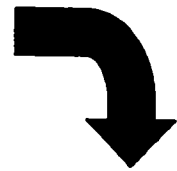


1. When something fails, we use it as a learning opportunity
2. Failure is not tolerated here
3. Individuals with ideas for service changes are supported
4. People are frightened to try new things
5. It is not safe here to try new ways of doing things
6. We don't have permission to make changes to practice
7. I am encouraged to be curious about practice
8. People here are encouraged to question the way that we do things
9. We are allowed to stand back and identify the need for practice change
10. We have lots of quality improvement initiatives, but little follow-through

[illegible]

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



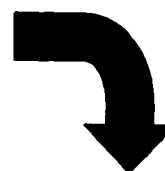
ORGANISATIONAL CHANGE: Leadership

	1	2	3	4	5
1. There is strong leadership to help us make changes in our practice	☐	☐	☐	☐	☐
2. We have access to support to deliver service changes	☐	☐	☐	☐	☐
3. Management here is receptive to ideas for changing service delivery	☐	☐	☐	☐	☐
4. If I had a good idea for service improvement, I know that I would be listened to	☐	☐	☐	☐	☐
5. If I had a good idea for service improvement, I'd know what to do to take it forward	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



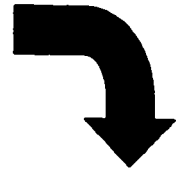
EVIDENCE BASED PRACTICE: Technical skills

	1	2	3	4	5
1. Critical appraisal skills training is available to those who want it	☐	☐	☐	☐	☐
2. We never use research without appraising it first	☐	☐	☐	☐	☐
3. I know where to get technical advice on evidence based practice	☐	☐	☐	☐	☐
4. I don't know who has the technical skills to help with evidence based practice	☐	☐	☐	☐	☐
5. There is training available in searching for research evidence	☐	☐	☐	☐	☐
6. Within our team, we have the skills to search for research evidence	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



EVIDENCE BASED PRACTICE: Access to information

	1	2	3	4	5
1. I have good access to information on research relevant to my job	☐	☐	☐	☐	☐
2. I have access to the internet at work	☐	☐	☐	☐	☐
3. My colleagues use research evidence as a source of information on best practice	☐	☐	☐	☐	☐
4. We receive useful information on our patient care outcomes	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

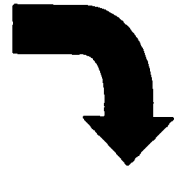
1= Strongly agree

2= Agree

3= Neither

4= Disagree

5= Disagree strongly

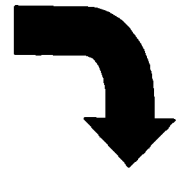


MULTI-DISCIPLINARY TEAMWORK: Common goals and objectives

	1	2	3	4	5
1. My colleagues and I share a common vision of service delivery	☐	☐	☐	☐	☐
2. The people I work with have shared goals	☐	☐	☐	☐	☐
3. I know what the goals for my service are	☐	☐	☐	☐	☐
4. The people I work with have clear service goals	☐	☐	☐	☐	☐
5. There is no clear vision of what it is that we are trying to achieve	☐	☐	☐	☐	☐
6. I don't feel part of a team	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

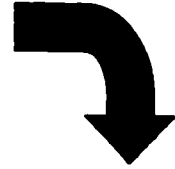
1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



	1	2	3	4	5
1. Cross-discipline working is the norm around here					
2. We are able to work around traditional hierarchies					
3. When it comes to patient care, no one profession is able to dominate					
4. Doctors unilaterally decide how services are run					
5. We have broken down the old tribalism between doctors, nurses and managers					
6. Some people's opinions carry more weight than others					
7. There is mutual respect for everyone's contribution					
8. Everyone has the same standing, regardless of professional background					
9. Colleagues are able to have robust discussions without generating bad feeling					
10. Colleagues value my expertise and skills					
11. People here seem to compete against each other					

[illegible]

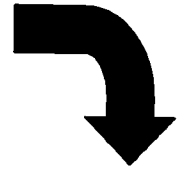
1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



1. I know what my colleagues expect of me
2. I have a good knowledge of the skills of those I work with
3. I have a good understanding of the different roles of my colleagues
4. My colleagues each have a specific role
5. I am clear about my role in service delivery
6. My colleagues don't seem to understand my role
7. People here don't have clear roles to play

[illegible]

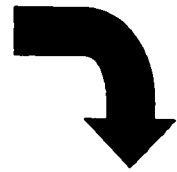
1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



1. It is safe to be open and honest with colleagues
2. I am open and honest with my colleagues
3. I feel able to raise clinical issues with my colleagues
4. I contribute fully to discussions with colleagues
5. Colleagues talk through their disagreements to achieve a consensus

1	2	3	4	5

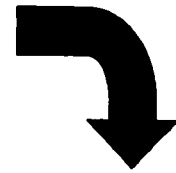
1= Strongly agree
2= Agree
3= Neither
4= Disagree
5= Disagree strongly



	1	2	3	4	5
1. We have shared standards of good clinical practice					
2. Clinical under-performance is addressed directly with the person concerned					
3. Clinical under-performance is ignored					
4. We have arrangements for detecting clinical performance problems early					
5. I could raise concerns about the poor performance of a colleague in confidence					
6. If I had concerns about the poor performance of a colleague, I would say so					
7. If there was a problem with my performance, I would receive help to develop					
8. If I raised concerns about a colleagues performance, it would be dealt with fairly					

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



APPRAISAL: reviews used developmentally

	1	2	3	4	5
1. I think staff appraisals are valuable	☐	☐	☐	☐	☐
2. Staff appraisals are an opportunity for personal development	☐	☐	☐	☐	☐
3. Staff appraisals are used to punish staff	☐	☐	☐	☐	☐
4. Staff appraisals are fair and balanced	☐	☐	☐	☐	☐
5. Staff appraisals are basically a way of catching people out	☐	☐	☐	☐	☐
6. I have to be very careful what I say in an appraisal	☐	☐	☐	☐	☐
7. My development needs are regularly assessed	☐	☐	☐	☐	☐
8. Appraisals are more about who gets a pay rise than development	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

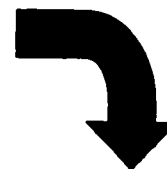
1= Strongly agree

2= Agree

3= Neither

4= Disagree

5= Disagree strongly



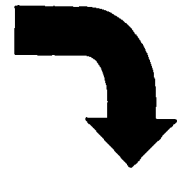
PROFESSIONAL DEVELOPMENT: Training and development needs

	1	2	3	4	5
1. My development needs are left unmet	☐	☐	☐	☐	☐
2. Training and development needs are never delivered on	☐	☐	☐	☐	☐
3. Training needs identified as part of development often remain unmet	☐	☐	☐	☐	☐
4. My career development needs are addressed alongside strategic needs of the service	☐	☐	☐	☐	☐
5. I've a clear understanding of how my career aims fit with the needs of the service	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

'Is the item a good indicator of the area of interest?'

- 1= Strongly agree
- 2= Agree
- 3= Neither
- 4= Disagree
- 5= Disagree strongly



PROFESSIONAL DEVELOPMENT: Using new skills

	1	2	3	4	5
1. Any new skills I develop are put to good use	☐	☐	☐	☐	☐
2. There are few opportunities to use new skills learned as part of development	☐	☐	☐	☐	☐
3. Once acquired, there is no recognition of new skills	☐	☐	☐	☐	☐
4. This organisation puts my skills to good use	☐	☐	☐	☐	☐
5. After training, there are few opportunities to practice and consolidate the learning	☐	☐	☐	☐	☐

Please make suggestions for re-wording of items, or offer alternative items, in the box below

Appendix 5: CGCQ Validation study tables

Appendix 5 lists respondent's opinions of the value of each of the draft indicator items included in the validation questionnaire. The tables also indicate what amendments were made to the item for the main development study questionnaire.

Original items where six or more respondents disagreed that the item was a good indicator are highlighted in **bold type**.
 Post-validation items incorporating minor changes from the original wording are highlighted in *italics*.
 For all tables, (n)= 31

ADVERSE EVENTS: Openness and encouragement to report adverse events

[Item a good indicator?: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
1. We are encouraged to admit our mistakes	16	10	3	1	0	1. We are encouraged to admit our mistakes
2. If I made a mistake I'd report it	23	4	3	1	0	Cut [confounds with personality type]
3. People here feel able to own-up to mistakes	14	8	5	3	0	2. <i>People own-up to their mistakes</i>
4. If someone makes a mistake, they try to cover it up	13	8	5	3	2	3. <i>People who make mistakes try to cover them up</i>
5. If I made a mistake I'd be open and honest about it	16	5	6	2	2	Cut [first person statement]
6. Only a fool would admit to making an error	6	8	6	5	6	4. <i>Only a fool would admit to making a mistake</i>
7. People are congratulated for coming forward about errors or mistakes	10	5	6	8	2	Cut [inability to discriminate]
8. People who identify adverse events are seen as just creating problems	14	8	4	2	3	5. <i>People who report adverse events are just creating problems</i>
9. Error reporting systems are basically a stick to beat clinicians with	5	7	6	4	8	6. Error reporting systems are basically a stick to beat clinicians with

ADVERSE EVENTS: Blame

[Item a good indicator?: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
12. When something goes wrong, there is a search for someone to blame	14	10	4	2	1	1. <i>When something goes wrong, there is an automatic assumption that 'someone is to blame'</i>
13. If I made a mistake I would be supported	17	6	5	2	1	2. <i>People who make mistakes are supported</i>
14. There is no support for staff involved in errors	18	5	3	4	1	3. <i>There is no support for staff involved in clinical incidents</i>
15. When things go wrong, there is an assumption that 'someone is to blame'	16	6	5	2	2	4. <i>When things go wrong, there is an assumption that 'someone is to blame'</i>
16. If I made a mistake, I'd be unjustly blamed for it	4	5	10	6	5	Cut [first person statement]
17. People involved in clinical incidents are made to feel guilty	15	8	4	3	1	5. <i>People involved in clinical incidents are made to feel guilty</i>
18. When there is an error, we look for failures in systems rather than blame individuals	18	7	4	1	1	6. <i>When there is an error, we look for failures in systems rather than blame individuals</i>
19. People involved in clinical incidents are made to feel like criminals	9	8	3	5	6	?Cut [loaded statement, close to item 5 above]
20. The emphasis is on how an incident happened, not who made the mistake	19	7	2	3	0	7. <i>The emphasis is on how a clinical incident happened, not who made the mistake</i>
21. We work in an atmosphere of blame	12	13	3	1	2	8. <i>We work in an atmosphere of blame</i>
22. If something goes wrong, someone will be scapegoated	13	8	4	3	3	Cut [weak statement]

ADVERSE EVENTS: Using a broad approach in investigations

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
8. The underlying causes of clinical incidents are simply ignored	16	6	4	2	3	1. The underlying causes of clinical incidents are simply ignored
9. When things go wrong, there is a thorough investigation to find the real causes	19	9	2	0	1	2. <i>When things go wrong, there is an investigation to find the real causes</i>
10. After an incident, there is a wide-ranging investigation	12	11	4	2	2	Cut ['wide-ranging' is too vague]
11. All of the circumstances surrounding a clinical incident are investigated	14	12	3	1	1	3. All of the circumstances surrounding a clinical incident are investigated
12. When things go wrong, we discuss the incident at length	14	6	6	3	2	Cut [too vague- formally or in the car park?]
13. We look beyond the immediate situation to find the real causes of clinical errors	16	9	2	2	2	4. We look beyond the immediate situation to find the real causes of clinical errors
14. There is pressure to 'solve' problems quickly rather than take time and do it properly	13	8	1	6	3	5. There is pressure to 'solve' problems quickly rather than take time and do it properly

ADVERSE EVENTS: Changes where necessary

[‘Item a good indicator?’: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
7. When things go wrong, we change practice to avoid it happening again	18	8	2	1	2	1. <i>When things go wrong, people change practice to avoid it happening again</i>
8. After an incident, we change our systems to stop the same conditions arising again	18	6	6	0	1	2. <i>After a clinical incident, changes are made to stop the same conditions arising again</i>
3. We've yet to change practice as a result of an investigation into an incident	16	6	4	4	1	3. <i>Recommendations made in the light of clinical incidents are just ignored</i>
5. Risk assessment processes are updated in the light of incidents	10	15	3	2	1	4. <i>Risk assessment processes are updated in the light of clinical incidents</i>

QUALITY IMPROVEMENT: Focus on organisational processes

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. We always check the effects on others of any quality improvements we make	9	13	6	2	1	1. <i>We consider the potential impact of our service improvements on colleagues across the organisation</i>
7. When making improvements, we try to get at the root causes of problems	12	12	1	4	2	Cut [conflates improvements and problems]
8. In solving one quality problem, we create others elsewhere	4	9	9	5	4	2. <i>In solving one quality problem, we create others elsewhere in the organisation</i>
9. The first we know of quality improvements elsewhere is when we feel their effects	8	10	8	4	1	3. <i>The first we know of quality improvements elsewhere in the organisation is when we feel their effects</i>
10. Often, improvements made elsewhere make our work more difficult	10	9	4	8	0	4. <i>Service improvements made elsewhere in the organisation make our work more difficult</i>

QUALITY IMPROVEMENT: Explicit focus on internal and external customers

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. We involve patients in our quality improvement activity	21	7	2	0	1	1. <i>We don't involve patients in our quality improvement activity</i>
7. We consult widely with colleagues when deciding priorities for quality improvement	12	12	6	1	0	2. We consult widely with colleagues when deciding priorities for quality improvement
8. Before making service improvements, we consult with patient representatives	13	9	6	3	0	3. <i>We consult with patient representatives before making service improvements</i>
9. We agree our quality improvement activity with those affected by changes	10	12	4	4	1	Cut ['woolly']
10. When improving services, we focus on service needs rather than patient needs	11	10	4	2	4	4. <i>Quality improvement activities tend to focus on service needs rather than patients needs</i>

QUALITY IMPROVEMENT: Use of cross-functional teams to address quality improvement

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. We work across teams to make improvements in whole systems of care	14	13	1	2	1	1. <i>We work together across teams to make quality improvements</i>
7. There is no cross-departmental work to improve quality	13	9	4	3	2	2. There is no cross-departmental quality improvement activity
8. We often work together across teams to improve our services	14	13	2	2	0	Cut [same as item 1 above]
9. People seem to fight their own corner rather than work together	10	14	0	3	4	3. People seem to fight their own corner rather than work together
10. We tend to improve our own areas, rather than work together to make big changes	14	6	6	4	1	4. We tend to improve our own areas, rather than work together to make big changes
						5. <i>Staff from different clinical teams work together on quality improvement</i>

QUALITY IMPROVEMENT: Investment in improvement activity

[‘Item a good indicator?’: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
8. Quality improvement is seen as an important activity	15	10	3	1	2	1. Quality improvement is seen as an important activity
9. We have protected time for quality improvement activity	15	10	4	2	0	2. We have protected time for quality improvement activity
3. There is no time to reflect on practice	18	7	3	2	1	3. <i>There is time to reflect on practice</i>
4. We continually try to improve the quality of our practice	9	9	6	5	2	Cut [too easy o say yes]
5. Clinicians are made to feel guilty about time away from direct patient contact	14	5	4	5	3	4. Clinicians are made to feel guilty about time away from direct patient contact
6. Immediate pressures are always more important than quality improvement	14	6	4	4	2	5. Immediate pressures are always more important than quality improvement
7. Long-term planning for quality improvement gets lost in the day-to-day	17	8	3	3	0	6. Long-term planning for quality improvement gets lost in the day-to-day

QUALITY IMPROVEMENT: Environment for improvement

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
8. Quality improvement is an important part of being a professional	10	13	5	2	1	Cut [not robust]
9. Quality improvement is imposed from above rather than built from below	11	11	5	4	0	1. Quality improvement is imposed from above rather than built from below
10. I feel involved in quality improvement- we are changing rather than being changed	19	10	1	0	1	2. <i>People feel involved in quality improvement- we are changing rather than being changed</i>
11. People are encouraged to lead on improvement in their areas of interest	18	8	4	1	0	3. People are encouraged to lead on improvement in their areas of interest
12. I am motivated to improve quality	16	10	4	1	0	4. <i>People are motivated to improve quality</i>
13. If I had an idea for improving quality, I'd know how to take it further	19	9	2	1	0	5. People are encouraged to develop ideas for improving quality
14. Our quality improvement activity is largely a response to external pressure	11	9	8	3	0	6. <i>Quality improvement activity is largely a response to external pressure</i>

QUALITY IMPROVEMENT: Use of problem solving techniques

[Item a good indicator?: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
5. We use specific quality improvement techniques	14	10	5	2	0	1. We use specific quality improvement techniques
6. We have no formal quality improvement processes	9	13	6	2	1	2. We have no formal quality improvement processes
7. I've been involved in developing care pathways	12	8	6	3	2	Cut [too limited]
8. When making quality improvements, we draw on a wide range of techniques	13	13	4	0	1	3. When making quality improvements, we draw on a wide range of techniques

QUALITY IMPROVEMENT: Planning for quality improvement

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. In our trust, quality improvement is about the latest fad	8	10	5	4	4	Cut [too ambiguous]
7. Around here, service improvements tend to be crisis-led	14	11	1	3	2	1. Around here, service improvements tend to be crisis-led
8. We have a clear system for selecting which innovations we will adopt	16	10	3	2	0	2. <i>There are clear systems for selecting which service innovations are adopted</i>
9. We base our improvement activity on an assessment of current problems	11	13	5	2	0	3. <i>Quality improvement activity is based on an assessment of current problems</i>
10. We use research evidence to inform our improvement activity	13	15	2	0	1	4. <i>Quality improvement activity is informed by research evidence</i>

RISK MANAGEMENT: A corporate approach

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
4. Clinical risk policies are shared throughout the organisation	20	10	1	0	0	1. Clinical risk policies are shared throughout the organisation
5. Each area handles clinical risk in its own way	13	10	3	3	3	Cut [bloated particular with item 2 below]
6. There is no common approach to risk management here	14	11	2	3	1	2. <i>There is no common approach to risk management</i>
4. We try to spot potential clinical incidents before they happen	16	10	5	0	0	3. <i>Potential clinical incidents are spotted before they happen</i>
5. We react to problems, rather than try to prevent them	16	10	2	2	1	4. We react to problems, rather than try to prevent them
7. We don't look at clinical risks in a systematic way	13	13	3	2	0	5. <i>Clinical risks are examined in a systematic way</i>
						6. We don't address the accidents waiting to happen

RISK MANAGEMENT: Assessment of risk

[Item a good indicator?: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. We systematically assess clinical risks	19	12	0	0	0	1. We systematically assess clinical risks
1. I'm conscious of the clinical risks involved in my work	13	11	5	2	0	Cut [first person]
2. We collect information on clinical risks	12	12	7	0	0	2. <i>Clinical risk information is used routinely to inform decisions</i>
3. We have no formal risk assessment systems	15	10	4	1	1	3. We have no formal risk assessment systems
4. We collect information on the clinical risks that matter most	8	10	8	4	1	4. <i>We don't collect information on the clinical risks that matter most</i>
						5. <i>Clinical risk information is used routinely</i>

RISK MANAGEMENT: Action on identified risks

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
1. When a clinical risk is identified, there is always action to address it	22	8	0	0	1	1. When a clinical risk is identified, there is always action to address it
2. Clinical risks are taken very seriously here	14	11	3	2	1	2. <i>Clinical risks are taken very seriously</i>
3. Identified clinical risks simply remain unaddressed	13	9	4	4	1	3. Identified clinical risks simply remain unaddressed
4. We react to incidents rather than plan to avoid them	13	10	5	3	0	4. We react to incidents rather than plan to avoid them
5. We don't address the accidents waiting to happen	15	8	1	5	2	Cut [move to 'corporate approach' above]
6. There is no sense of urgency in managing clinical risks	13	9	4	4	1	5. There is no sense of urgency in addressing clinical risks
						6. <i>Risk management is largely a paper exercise</i>

ORGANISATIONAL LEARNING: Investment in learning

[‘Item a good indicator?’: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
7. Time spent reflecting on practice is seen as 'not real work'	14	9	4	3	1	1. Time spent reflecting on practice is seen as 'not real work'
8. There is no time to get together to share ideas	11	14	4	2	0	2. There is no time to get together to share ideas
9. We are encouraged to spend time sharing ideas with others	15	10	4	2	0	3. <i>People spend time sharing ideas with others</i>
10. We are encouraged to devote time to disseminating our good practice	11	16	4	0	0	4. <i>People devote time to disseminating good practice</i>
11. Reflection on practice is accepted as a valuable part of being a professional	15	12	3	0	1	Cut [too easy to say 'yes']
12. Sharing quality improvements is seen as an essential part of good practice	15	12	2	2	0	5. Sharing quality improvements is an essential part of good practice

ORGANISATIONAL LEARNING: Sharing learning with others

[‘Item a good indicator?’: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
11. I don't know about good practice taking place in other parts of the organisation	19	6	2	3	1	1. <i>People don't know about good practice taking place in other parts of the organisation</i>
12. We share practice issues with others in different parts of the organisation	13	13	2	3	0	2. We share practice issues with others in different parts of the organisation
13. People tend to hold on to information for their own advantage	10	8	5	5	3	Cut [not related to area]
14. In this organisation, good practice stays isolated in pockets	12	15	0	3	1	3. <i>Good practice stays isolated in pockets</i>
15. When we learn something useful, we share it across the whole trust	20	7	3	0	1	4. When we learn something useful, we share it across the whole trust
16. Around here, teams from different parts of the organisation share their good practice	16	9	3	2	1	5. <i>Teams from different parts of the organisation share their good practice</i>
17. We share good practice ideas with others outside the organisation	17	10	2	2	0	6. <i>Good practice ideas are shared with others outside the organisation</i>
18. There is support for sharing insights on practice across the organisation	11	10	6	3	1	Cut [Vague: from whom?]
19. We are keen to learn from others outside our organisation	16	9	4	0	2	7. <i>We don't seem able to learn from others outside our organisation</i>
20. We have a 'not invented here' attitude to new ideas	5	7	9	6	4	Cut [confusing]

ORGANISATIONAL LEARNING: Motivation to change

[‘Item a good indicator?’: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
2. I can't see the benefit of making changes to services	10	6	5	5	5	Cut [Too obvious]
10. People here are free to question 'custom and practice'	16	12	3	0	0	1. <i>People are free to question 'custom and practice'</i>
11. Most people here want to work in the same way that they always have done	11	13	4	1	2	2. <i>People want to work in the same way that they always have done</i>
12. People here are highly motivated to make changes to clinical practice	16	11	3	1	0	3. <i>People are highly motivated to make changes to clinical practice</i>
13. People here are prepared to try new ways of doing things	21	4	3	0	0	4. <i>People are prepared to try new ways of doing things</i>
14. There is a great deal of inertia which stops most changes to practice	11	9	4	5	2	Cut [Unclear]
15. It is not my place to suggest changes to services	11	11	1	5	3	Cut [First person]
16. We tend to do things as they have always been done	13	11	2	2	3	5. <i>People tend to do things as they have always been done</i>
17. We are forced into making service changes, rather than encouraged to make them	12	10	7	0	2	6. <i>People are forced into making service changes, rather than encouraged to make them</i>

ORGANISATIONAL CHANGE: Support

[‘Item a good indicator?’: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
11. When something fails, it is used as a learning opportunity	10	9	2	1	0	1. When something fails, it is used as a learning opportunity
12. Failure is not tolerated here	16	6	4	2	3	2. <i>Failure is not tolerated</i>
13. Individuals with ideas for service changes are supported	15	13	3	0	0	3. People with good ideas for service changes are often ignored
14. People are frightened to try new things	11	13	2	4	1	Cut [repeat of item 4 on 'motivation to change']
15. It is not safe here to try new ways of doing things	9	9	6	7	2	Cut [confusing]
16. We don't have permission to make changes to practice	12	12	0	6	1	Cut [Confusing]
17. I am encouraged to be curious about practice	10	15	3	1	3	Cut [First person]
18. People here are encouraged to question the way that we do things	14	15	1	0	1	4. <i>People are encouraged to question the way that things are done</i>
19. We are allowed to stand back and identify the need for practice change	11	14	2	4	0	5. <i>People are given time to stand back and identify the need for practice change</i>
20. We have lots of quality improvement initiatives, but little follow-through	11	12	7	1	0	6. <i>There are lots of quality improvement initiatives, but little real change</i>

ORGANISATIONAL CHANGE: Leadership

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. There is strong leadership to help us make changes in our practice	17	9	2	2	1	1. There is strong leadership to help people make changes in their practice
7. We have access to support to deliver service changes	17	6	5	2	1	2. <i>There is no support to deliver service changes</i>
8. Management here is receptive to ideas for changing service delivery	16	11	3	1	0	3. <i>Management is receptive to ideas for changing service delivery</i>
9. If I had a good idea for service improvement, I know that I would be listened to	16	10	4	1	0	4. <i>People with good ideas for service development are listened to</i>
10. If I had a good idea for service improvement, I'd know what to do to take it forward	15	12	2	1	1	5. <i>Good service improvement ideas are taken forward</i>

EVIDENCE BASED PRACTICE: Technical skills

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
7. Critical appraisal skills training is available to those who want it	13	13	5	0	1	1. Critical appraisal skills training is available to those who want it
8. We never use research without appraising it first	15	8	2	5	1	2. <i>Research is appraised before use</i>
9. I know where to get technical advice on evidence based practice	15	15	0	1	0	3. <i>There is no advice available on Evidence Based Practice</i>
10. I don't know who has the technical skills to help with evidence based practice	11	14	3	3	0	4. <i>Technical help with Evidence Based Practice is available</i>
11. There is training available in searching for research evidence	18	9	4	0	0	5. There is no training available in searching for research evidence
12. Within our team, we have the skills to search for research evidence	13	10	6	2	0	6. <i>Clinical teams don't have the skills to search for research evidence</i>

EVIDENCE BASED PRACTICE: Access to information

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
5. I have good access to information on research relevant to my job	21	7	1	1	2	1. <i>People have no access to research relevant to their practice</i>
6. I have access to the internet at work	15	13	1	1	1	2. <i>Those that need it have access to the internet at work</i>
7. My colleagues use research evidence as a source of information on best practice	14	11	4	2	0	3. <i>People use research evidence as a source of information on best practice</i>
8. We receive useful information on our patient care outcomes	17	9	3	1	1	4. <i>People receive useful information on patient care outcomes</i>

MULTI-DISCIPLINARY TEAMS: Common goals and objectives

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
7. My colleagues and I share a common vision of service delivery	14	10	4	2	1	1. <i>People share a common vision of service delivery</i>
8. The people I work with have shared goals	12	13	3	2	1	2. <i>People don't seem to have shared service goals</i>
9. I know what the goals for my service are	16	9	2	3	1	3. <i>People know their service goals</i>
10. The people I work with have clear service goals	16	11	2	2	0	4. <i>Service goals are unclear</i>
11. There is no clear vision of what it is that we are trying to achieve	15	11	1	2	2	5. <i>There is no clear vision of what it is the organisation is trying to achieve</i>
12. I don't feel part of a team	14	6	3	3	2	Cut [First person]

MULTI-DISCIPLINARY TEAMWORK: Mutual respect

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
12. Cross-discipline working is the norm around here	10	13	6	1	1	1. <i>Cross-discipline working is the norm</i>
13. We are able to work around traditional hierarchies	8	14	5	3	1	2. <i>People work around traditional hierarchies</i>
14. When it comes to patient care, no one profession is able to dominate	13	10	2	6	0	3. When it comes to patient care, no one profession is able to dominate
15. Doctors unilaterally decide how services are run	14	10	1	4	2	4. <i>Doctors generally decide how services are run</i>
16. We have broken down the old tribalism between doctors, nurses and managers	15	11	4	0	1	5. <i>The old tribalisms between doctors, nurses and managers still remain</i>
17. Some people's opinions carry more weight than others	15	11	2	2	1	6. Some people's opinions carry more weight than others
18. There is mutual respect for everyone's contribution	15	15	1	0	0	7. There is mutual respect for everyone's contribution
19. Everyone has the same standing, regardless of professional background	12	12	3	3	0	8. Everyone has the same standing, regardless of professional background
20. Colleagues are able to have robust discussions without generating bad feeling	11	12	4	2	2	Cut [move to 'participation' section]
21. Colleagues value my expertise and skills	9	15	3	4	0	9. <i>Colleagues don't seem to value each others expertise and skills</i>
22. People here seem to compete against each other	9	11	4	4	3	Cut

MULTI-DISCIPLINARY TEAMWORK: Clear roles

[Item a good indicator?: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
8. I know what my colleagues expect of me	17	11	2	1	0	1. <i>People don't know what their colleagues expect of them</i>
9. I have a good knowledge of the skills of those I work with	14	12	3	2	0	2. <i>People have a good knowledge of the skills of their colleagues</i>
10. I have a good understanding of the different roles of my colleagues	17	12	1	1	0	3. <i>People have a good understanding of the different roles of their colleagues</i>
11. My colleagues each have a specific role	10	14	1	5	1	Cut [Confusing]
12. I am clear about my role in service delivery	15	9	2	3	1	4. <i>People are clear about their role in service delivery</i>
13. My colleagues don't seem to understand my role	13	8	5	4	1	5. <i>Colleagues don't seem to understand each others role's</i>
14. People here don't have clear roles to play	14	10	2	1	4	Cut [confusing]

MULTI-DISCIPLINARY TEAMWORK: Participation

[Item a good indicator?: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. It is safe to be open and honest with colleagues	15	7	5	4	0	9. It is unsafe to be open and honest with colleagues
7. I am open and honest with my colleagues	17	9	3	1	1	10. Colleagues are dishonest with each other
8. I feel able to raise clinical issues with my colleagues	20	7	2	2	0	11. People raise clinical issues with their colleagues
9. I contribute fully to discussions with colleagues	13	14	2	2	0	4. People contribute fully to discussions with colleagues
10. Colleagues talk through their disagreements to achieve a consensus	17	10	1	2	1	5. Colleagues talk through their disagreements
6. Colleagues are able to have robust discussions without generating bad feeling	11	12	4	2	2	6. Colleagues are able to have robust discussions without generating bad feeling

APPRAISAL: Managing poor clinical performance

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
1. We have shared standards of good clinical practice	15	8	3	3	0	Cut [Confusing]
2. Clinical under-performance is addressed directly with the person concerned	16	12	2	1	0	1. <i>Poor clinical performance is addressed with the person concerned</i>
3. Clinical under-performance is ignored	16	9	2	1	3	2. <i>Poor clinical performance is ignored</i>
4. We have arrangements for detecting clinical performance problems early	17	11	3	0	0	3. <i>Poor clinical performance goes largely undetected</i>
5. I could raise concerns about the poor performance of a colleague in confidence	19	9	3	0	0	4. <i>People can raise concerns about a colleague's poor clinical performance in confidence</i>
6. If I had concerns about the poor performance of a colleague, I would say so	16	12	3	0	0	Cut [Confounded by personality]
7. If there was a problem with my performance, I would receive help to develop	16	11	3	1	0	5. <i>People with performance problems don't receive any help to develop</i>
8. If I raised concerns about a colleagues performance, it would be dealt with fairly	13	14	3	1	0	6. <i>Concerns about poor clinical performance are dealt with fairly</i>

APPRAISAL: Reviews used developmentally

[‘Item a good indicator?’: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
9. I think staff appraisals are valuable	17	12	1	1	0	1. <i>Staff appraisals are a waste of time</i>
10. Staff appraisals are an opportunity for personal development	18	12	1	0	0	2. Staff appraisals are an opportunity for personal development
11. Staff appraisals are used to punish staff	9	6	5	5	6	3. Staff appraisals are used to punish staff
12. Staff appraisals are fair and balanced	8	14	7	2	0	4. Staff appraisals are fair and balanced
13. Staff appraisals are basically a way of catching people out	8	8	6	3	6	Cut
14. I have to be very careful what I say in an appraisal	15	6	6	1	3	5. <i>People find it difficult to be honest in appraisals</i>
15. My development needs are regularly assessed	14	13	2	1	1	6. <i>Development needs are regularly assessed</i>
16. Appraisals are more about who gets a pay rise than development	3	10	3	5	7	Cut

APPRAISAL: Training and development needs

[Item a good indicator?: 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. My development needs are left unmet	15	9	2	1	4	1. <i>My development needs are met</i>
7. Training and development needs are never delivered on	11	12	3	3	2	2. <i>Training and development needs are never addressed</i>
8. Training needs identified as part of development often remain unmet	11	11	5	2	2	3. <i>Training needs are ignored</i>
9. My career development needs are addressed alongside strategic needs of the service	16	10	4	0	1	4. <i>career development needs are addressed alongside strategic needs of the service</i>
10. I've a clear understanding of how my career aims fit with the needs of the service	17	13	0	0	1	Cut [first person]
						5. <i>Appraisals do not identify the real needs of staff</i>

APPRAISAL: Using new skills

['Item a good indicator?': 1= Strongly agree, 2= Agree, 3= Neither, 4= Disagree, 5= Strongly disagree]

Original item	Item a good indicator?					Post-validation item
	1	2	3	4	5	
6. Any new skills I develop are put to good use	18	10	2	1	0	1. <i>New skills developed in training are put to good use</i>
7. There are few opportunities to use new skills learned as part of development	14	10	2	3	2	2. There are few opportunities to use new skills learned as part of development
8. Once acquired, there is no recognition of new skills	10	10	4	4	3	Cut
9. This organisation puts my skills to good use	16	11	2	1	1	3. <i>People's skills are put to good use</i>
10. After training, there are few opportunities to practice and consolidate the learning	13	11	4	2	1	4. <i>After training, there are few opportunities to practice and consolidate learning</i>

Appendix 6: The CGCQ schedule used in the development study sample

This appendix contains a complete copy of the questionnaire sent to a sample of 1140 health care staff between December 2001 and May 2002.



**THE UNIVERSITY
OF BIRMINGHAM**

**The Clinical Governance Climate Questionnaire
(CGCQ)**

**A measure of staff perceptions of progress
in clinical governance**

© Tim Freeman HSMC 2001

Our aim in this questionnaire is to obtain your views about the climate of clinical governance in your organisation.

Clinical governance is a prominent issue in health policy circles, and its impact ultimately depends on how people respond. There are very few studies, however, that have examined what its effects are upon the daily activities of those most affected. We therefore seek your opinion on a broad range of factors within your organisation, such as: managing clinical risks; quality improvement; professional development and sharing good practice.

There are currently no measures of the impact of clinical governance on people's working environment. We are currently undertaking development work for such a measure here at HSMC, and are seeking your help to reduce a large number of potential indicators into much smaller number of reliable indicators. We realise the demanding nature of the task we are asking you to do, and would be most grateful if you would complete and return the questionnaire to us here at HSMC.

Some additional points

1. Funded by the NHS Executive, the study is part of a linked series of research activities into clinical governance.
2. We give an assurance that individuals responding to this questionnaire will not be identifiable in any reports of the findings. Your responses will be held in complete confidence by the research team at HSMC, the University of Birmingham.
3. The statements contained in the questionnaire do not necessarily represent the views of HSMC. You may regard some of the statements as polemical. They are included in the survey so that you can express your views about them.
4. **There are no right or wrong answers to the questions. We are simply interested in your views.**

Thank you for giving us some of your valuable time.

Personal Details

We would like to know a few personal details. We will use this information when we analyse your responses to the rest of the questionnaire.

1. The name of the trust where you work _____

2. Length of employment with current employer

Less than 1 year	☐ 1
1 - 4 years	☐ 2
5 - 9 years	☐ 3
10 - 14 years	☐ 4
15 years or more	☐ 5

3. Your professional or occupational role

Medical (surgery)	☐ 1	Managerial	☐ 6
Medical (medicine)	☐ 2	Quality/ Research	☐ 7
GP	☐ 3	Scientific/ Laboratory	☐ 8
Nursing/ HV/ Midwifery	☐ 4	Administrative/ Clerical	☐ 9
PAMs / Therapies/ Psychology	☐ 5	Technical/ Crafts	☐ 10

Other (please specify) ☐11 _____

4. Your age in years

Years

5. Do you spend more than 5% of your time on managerial activities?

Yes	☐ 1
No	☐ 2

6. Which of the following best describes your current grade?
(choose one only)

Medical

Consultant / Specialist	☐ 1
Registrar	☐ 2
SHO/ HO	☐ 3
GP	☐ 4
Other (please specify)	☐ 5
-

Nursing and allied

Grade 'H' or above	☐ 6
Grades 'F' to 'G'	☐ 7
Grades C to E	☐ 8
Unqualified / Health Care Assistant	☐ 9
Other (please specify)	☐ 10
-

Managerial

Executive director	☐ 11
Divisional director	☐ 12
Directorate/General /Deputy departmental manager	☐ 13
Deputy	☐ 14
Other (please specify)	☐ 15
-

PAMs / Psychology

Senior therapist/ Senior psychologist	☐ 16
Therapist/ Psychologist	☐ 17
Other (please specify)	☐ 18
-

Quality/ Administrative/ Clerical/ Technical

Facilitator / Clinical audit/ Research staff	☐ 19
Scientific officer	☐ 20
Administrative / Clerical	☐ 21
Technical / craft	☐ 22

Other (please specify)

☐ 23

Other (please specify)

☐ 24

Please identify your highest qualification in each of the following areas:

7 (a) Medicine

None	☐ 1
Undergraduate degree	☐ 2
Postgraduate certificate or diploma	☐ 3
Postgraduate degree	☐ 4
Fellowship	☐ 5
Other (please specify)	☐ 6
-

7 (b) Nursing

None	☐ 1
Basic registration	☐ 2
Post-basic certificate or diploma	☐ 3
Undergraduate degree	☐ 4
Postgraduate certificate or diploma	☐ 5
Postgraduate degree	☐ 6
Other (please specify)	☐ 7
-

7 (c) Management

None	☐ 1
Short course	☐ 2
Certificate or diploma	☐ 3
Undergraduate degree	☐ 4
Postgraduate certificate or diploma	☐ 5
Postgraduate degree	☐ 6
Other (please specify)	☐ 7
-

7 (d) Other disciplines

None	☐ 1
Certificate (e.g. ONC)	☐ 2
Diploma(e.g. HND)	☐ 3
Undergraduate degree	☐ 4
Postgraduate certificate or diploma	☐ 5
Postgraduate degree	☐ 6

Other (please specify)

☐7

8. Please identify which of the following directorates you work in
(choose one only)

Mental Health & Learning Disabilities	☐ 1
Rehabilitation	☐ 2
Elderly care	☐ 3
Community and children's	☐ 4
Headquarters	☐ 5
Facilities	☐ 6

Thank you for this information

Now turn you attention to the next section, on your experience of the place where you work

Your experience of the place where you work

We would like to know what it is like to work in your organisation

Thinking of the part of the organisation where you work, indicate whether you agree or disagree with each of the following statements. Please rate all the statements on the 5-point scale by circling the number that best represents your opinion.

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

<i>In the part of the organisation where I work...</i>		<i>(Circle)</i>				
1	When things go wrong, there is an automatic assumption that 'someone is to blame'	1	2	3	4	5
2	We don't seem able to learn from others outside our organisation	1	2	3	4	5
3	We tend to improve our own clinical areas, rather than work together to make big changes	1	2	3	4	5
4	People tend to do things as they have always been done	1	2	3	4	5
5	When something goes wrong, there is a search for someone to blame	1	2	3	4	5
6	We look beyond immediate situations to find the real causes of clinical errors	1	2	3	4	5
7	Colleagues talk through their disagreements	1	2	3	4	5
8	People are encouraged to question the way that things are done	1	2	3	4	5
9	When things go wrong, there is an investigation to find the real causes	1	2	3	4	5
10	Good practice ideas are shared with others outside the organisation	1	2	3	4	5
11	There is no support for staff involved in clinical incidents	1	2	3	4	5
12	Good service improvement ideas are taken forward	1	2	3	4	5
13	All of the circumstances surrounding a clinical incident are investigated	1	2	3	4	5
14	There is no advice available on evidence based practice	1	2	3	4	5
15	People seem to fight their own corner rather than work together	1	2	3	4	5
16	People with good ideas for service changes are often ignored	1	2	3	4	5
17	We have no formal quality improvement processes	1	2	3	4	5
18	Colleagues don't seem to value each others expertise and skills	1	2	3	4	5
19	People are free to question 'custom and practice'	1	2	3	4	5
20	Sharing quality improvements is an essential part of good practice	1	2	3	4	5
21	People are given time to stand back and identify the need for practice change	1	2	3	4	5
22	Error reporting systems are basically a stick to beat clinicians with	1	2	3	4	5
23	Cross-discipline working is the norm	1	2	3	4	5
24	People use research evidence as a source of information on best practice	1	2	3	4	5
25	Critical appraisal skills training is available to those who want it	1	2	3	4	5
26	There is strong leadership to help people make changes to their practice	1	2	3	4	5
27	People involved in clinical incidents are made to feel guilty	1	2	3	4	5
28	Career development needs are addressed alongside strategic needs of the service	1	2	3	4	5
29	Staff appraisals are used to punish staff	1	2	3	4	5
30	People have a good knowledge of the skills of their colleagues	1	2	3	4	5
31	We have protected time for quality improvement activity	1	2	3	4	5
32	It is unsafe to be open and honest with colleagues	1	2	3	4	5
33	Potential clinical incidents are spotted before they happen	1	2	3	4	5
34	The emphasis is on how an incident happened, not who made the mistake	1	2	3	4	5
35	People who make mistakes try to cover them up	1	2	3	4	5
36	Training needs are ignored	1	2	3	4	5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

	<i>In the part of the organisation where I work...</i>	<i>(Circle)</i>
37	We have no formal risk assessment systems	1 2 3 4 5
38	People own-up to their mistakes	1 2 3 4 5
39	Training and development needs are never addressed	1 2 3 4 5
40	There is no cross-departmental work to improve quality	1 2 3 4 5
41	Poor clinical performance is ignored	1 2 3 4 5
42	People who make mistakes are supported	1 2 3 4 5
43	Poor clinical performance goes largely undetected	1 2 3 4 5
44	We work together across teams to make quality improvements	1 2 3 4 5
45	People who report adverse events are just creating problems	1 2 3 4 5
46	Staff appraisals are fair and balanced	1 2 3 4 5
47	People contribute fully to discussions with colleagues	1 2 3 4 5
48	The old tribalisms between doctors, nurses and managers remain	1 2 3 4 5
49	Service goals are unclear	1 2 3 4 5
50	We collect information on clinical risks	1 2 3 4 5
51	People with good service improvement ideas are listened to	1 2 3 4 5
52	Technical help with evidence based practice is available	1 2 3 4 5
53	When there is an error, we look for failures in systems rather than blame individuals	1 2 3 4 5
54	People have a good understanding of the different roles of their colleagues	1 2 3 4 5
55	Appraisal does not identify the real development needs of staff	1 2 3 4 5
56	Long-term planning for quality improvement gets lost in the day-to-day	1 2 3 4 5
57	Professional development needs are met	1 2 3 4 5
58	People receive useful information on patient care outcomes	1 2 3 4 5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

	<i>In the part of the organisation where I work...</i>	<i>(Circle)</i>
59	Colleagues are dishonest with each other	1 2 3 4 5
60	We are encouraged to admit our mistakes	1 2 3 4 5
61	New skills developed in training are put to good use	1 2 3 4 5
62	Failure is not tolerated	1 2 3 4 5
63	When making quality improvements, we draw on a wide range of techniques	1 2 3 4 5
64	Good practice stays isolated in pockets	1 2 3 4 5
65	People feel involved in quality improvement- we are changing rather than being changed	1 2 3 4 5
66	Identified clinical risks simply remain unaddressed	1 2 3 4 5
67	Clinical risks are examined in a systematic way	1 2 3 4 5
68	People are encouraged to lead on improvement in their areas of interest	1 2 3 4 5
69	Quality improvement activities tend to focus on service needs rather than patient needs	1 2 3 4 5
70	Time spent reflecting on practice is seen as 'not real work'	1 2 3 4 5
71	People share practice issues with others in different parts of the organisation	1 2 3 4 5
72	Clinical risks are taken very seriously	1 2 3 4 5
73	Staff from different clinical teams work together on quality improvement	1 2 3 4 5
74	People find it difficult to be honest in an appraisal	1 2 3 4 5
75	People can raise concerns about a colleague's poor clinical performance in confidence	1 2 3 4 5
76	Clinical teams don't have the skills to search for research evidence	1 2 3 4 5
77	There is no support to deliver service changes	1 2 3 4 5
78	When it comes to patient care, no one profession is able to dominate	1 2 3 4 5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

	<i>In the part of the organisation where I work...</i>	<i>(Circle)</i>
79	The underlying causes of clinical incidents are simply ignored	1 2 3 4 5
80	There is no clear vision of what it is that the organisation is trying to achieve	1 2 3 4 5
81	We work in an atmosphere of blame	1 2 3 4 5
82	Doctors generally decide how services are run	1 2 3 4 5
83	We don't collect information on the clinical risks that matter most	1 2 3 4 5
84	After a clinical incident, changes are made to stop the same conditions arising again	1 2 3 4 5
85	People raise clinical issues with their colleagues	1 2 3 4 5
86	There is no training available in searching for research evidence	1 2 3 4 5
87	There are lots of quality improvement initiatives, but little real change	1 2 3 4 5
88	Quality improvement activity is based on an assessment of current problems	1 2 3 4 5
89	People spend time sharing ideas with others	1 2 3 4 5
90	There is no common approach to risk management	1 2 3 4 5
91	Teams from different parts of the organisation share their good practice	1 2 3 4 5
92	We react to incidents rather than plan to avoid them	1 2 3 4 5
93	People devote time to disseminating good practice	1 2 3 4 5
94	The first we know of quality improvements elsewhere in the organisation is when we feel their effects	1 2 3 4 5
95	Staff appraisals are an opportunity for personal development	1 2 3 4 5
96	Service improvements tend to be crisis-led	1 2 3 4 5
97	When a clinical risk is identified, there is always action to address it	1 2 3 4 5
98	Concerns about poor clinical performance are dealt with fairly	1 2 3 4 5
99	Quality improvement is imposed from above rather than built from below	1 2 3 4 5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

	<i>In the part of the organisation where I work...</i>	<i>(Circle)</i>
100	Some people's opinions carry more weight than others	1 2 3 4 5
101	We systematically assess clinical risks	1 2 3 4 5
102	Research is appraised before use	1 2 3 4 5
103	There are clear systems for selecting which service innovations are adopted	1 2 3 4 5
104	There is no time to get together to share ideas	1 2 3 4 5
105	Colleagues are able to have robust discussions without generating bad feeling	1 2 3 4 5
106	People share a common vision of service delivery	1 2 3 4 5
107	There is no sense of urgency in addressing clinical risks	1 2 3 4 5
108	There is mutual respect for everyone's contribution	1 2 3 4 5
109	There is pressure to 'solve' problems quickly rather than take time and do it properly	1 2 3 4 5
110	We don't address the accidents waiting to happen	1 2 3 4 5
111	When things go wrong, people change practice to avoid it happening again	1 2 3 4 5
112	People are encouraged to develop ideas for improving quality	1 2 3 4 5
113	People are prepared to try new ways of doing things	1 2 3 4 5
114	Recommendations made in the light of clinical incidents are just ignored	1 2 3 4 5
115	When we learn something useful, we share it across the whole trust	1 2 3 4 5
116	Quality improvement activity is informed by research evidence	1 2 3 4 5
117	Clinical risk policies are shared throughout the organisation	1 2 3 4 5
118	Clinical risk information is used routinely to inform decisions	1 2 3 4 5
119	Service improvements made elsewhere in the organisation make our work more difficult	1 2 3 4 5
120	Only a fool would admit to making a mistake	1 2 3 4 5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

	<i>In the part of the organisation where I work...</i>	<i>(Circle)</i>
121	Clinicians are made to feel guilty about time away from direct patient contact	1 2 3 4 5
122	People want to work in the same way that they always have done	1 2 3 4 5
123	Risk management is largely a paper exercise	1 2 3 4 5
124	People don't know what their colleagues expect of them	1 2 3 4 5
125	People don't seem to have shared service goals	1 2 3 4 5
126	There is time to reflect on practice	1 2 3 4 5
127	Staff appraisals are a waste of time	1 2 3 4 5
128	People don't know about good practice taking place in other parts of the organisation	1 2 3 4 5
129	Development needs are regularly assessed	1 2 3 4 5
130	Management is receptive to ideas for changing service delivery	1 2 3 4 5
131	People know their service goals	1 2 3 4 5
132	In solving one quality problem, we create others elsewhere in the organisation	1 2 3 4 5
133	We consult widely with colleagues when deciding quality improvement priorities	1 2 3 4 5
134	People are clear about their role in service delivery	1 2 3 4 5
135	People work around traditional hierarchies	1 2 3 4 5
136	Those that need it have access to the internet at work	1 2 3 4 5
137	We don't involve patients in our quality improvement activity	1 2 3 4 5
138	Immediate pressures are always more important than quality improvement	1 2 3 4 5
139	Quality improvement activity is largely a response to external pressure	1 2 3 4 5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

	<i>In the part of the organisation where I work...</i>	<i>(Circle)</i>
140	We consider the potential impact of our service improvements on colleagues across the organisation	1 2 3 4 5
141	After training, there are few opportunities to practice and consolidate learning	1 2 3 4 5
142	Risk assessment processes are updated in the light of clinical incidents	1 2 3 4 5
143	People with performance problems don't receive any help to develop	1 2 3 4 5
144	Quality improvement is seen as an important activity	1 2 3 4 5
145	People's skills are put to good use	1 2 3 4 5
146	Colleagues don't seem to understand each others role's	1 2 3 4 5
147	When something fails, it is used as a learning opportunity	1 2 3 4 5
148	We consult with patient representatives before making service improvements	1 2 3 4 5
149	People have no access to research relevant to their practice	1 2 3 4 5
150	We use specific quality improvement techniques	1 2 3 4 5
151	Everyone has the same standing, regardless of professional background	1 2 3 4 5
152	We react to problems, rather than try to prevent them	1 2 3 4 5
153	People are motivated to improve quality	1 2 3 4 5
154	There are few opportunities to use new skills learned as part of development	1 2 3 4 5
155	People are forced into making service changes, rather than encouraged to make them	1 2 3 4 5
156	People are highly motivated to make changes to clinical practice	1 2 3 4 5
157	Poor clinical performance is addressed directly with the person concerned	1 2 3 4 5

How you feel about your work

We would like to know how you currently feel about your work

Please rate all the items on the 5-point scale by circling the number that best represents your opinion.

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

(Circle)

1	I feel emotionally drained from my work	1	2	3	4	5
2	I feel used up at the end of the workday	1	2	3	4	5
3	I feel tired when I get up in the morning and have to face another day on the job	1	2	3	4	5
4	Working all day is really a strain for me	1	2	3	4	5
5	I can effectively solve the problems that arise in my work	1	2	3	4	5
6	I feel burned out from my work	1	2	3	4	5
7	I feel I am making an effective contribution to what this organisation does	1	2	3	4	5
8	I have become less interested in my work since I started this job	1	2	3	4	5
9	I have become less enthusiastic about my work	1	2	3	4	5
10	In my opinion, I am good at my job	1	2	3	4	5
11	I feel exhilarated when I accomplish something at work	1	2	3	4	5
12	I have accomplished many worthwhile things in this job	1	2	3	4	5
13	I just want to do my job and not be bothered	1	2	3	4	5
14	I have become more cynical about whether my work contributes anything	1	2	3	4	5
15	I doubt the significance of my work	1	2	3	4	5
16	At my work, I feel confident that I am effective at getting things done	1	2	3	4	5

What your organisation is like

These questions relate to the type of organisation that your workplace is most like

Each of these items contains four descriptions of organisations. Please distribute 100 points among the four descriptions depending on how *similar* the description is to your organisation. None of the descriptions is any better than the others; they are just different. *For each question, please use all 100 points.*

FOR EXAMPLE: In question 1, if organisation A seems very similar to mine, B somewhat similar, and C and D do not seem similar at all, I might give 70 points to A and the remaining 30 points to B.

1. Organisation characteristics (please distribute 100 points):

- A is a very *personal* place. It is like an extended family. People seem to share a lot of themselves _____ %
- B is a very *dynamic* place. People are willing to stick their necks out and take risks _____ %
- C is a very *formalised and structured* place. Bureaucratic procedures govern what people do _____ %
- D is very *production oriented*. The major concern is getting the job done _____ %

2. Organisation leader (please distribute 100 points):

- The Chief Executive of A is generally considered a *mentor, sage or parental figure* _____ %
- The Chief Executive of B is generally considered an *entrepreneur, innovator or risk-taker* _____ %
- The Chief Executive of C is generally considered a *co-ordinator, organiser or administrator* _____ %
- The Chief Executive of D is generally considered a *producer, technician or hard-driver* _____ %

3. The organisations 'glue' (please distribute 100 points):

- The glue that holds A together is *loyalty and tradition* _____ %
- The glue that holds B together is a *commitment to innovation and development* _____ %
- The glue that holds C together is adherence to *formal rules and policies* _____ %
- The glue that holds D together is an emphasis on *tasks and goal accomplishment* _____ %

4. Organisational emphasis (please distribute 100 points):

- A emphasises *human resources*. High cohesion and morale are important _____ %
- B emphasises *growth and new resources*. Readiness to meet new challenges is important _____ %
- C emphasises *permanence and stability*. Measurable goals are important _____ %
- D emphasises *competitive actions and achievement*. Efficient, smooth operations are important _____ %

Thank you for completing this questionnaire

We would be most grateful if you would return it to us here at HSMC in the pre-paid envelope provided.

If you would like further information on the study, then please do not hesitate to contact me by e-mail, telephone or letter:

Mr. T. Freeman
Lecturer
Health Services Management Centre
University of Birmingham
Park House
40 Edgbaston Park Road
Birmingham B15 2RT

Tel: 

e-mail: 

Once again, many thanks for your time and co-operation; it is very much appreciated.

Appendix 7: The 60-item short form CGCQ

This appendix reproduced the 60-item short form of the CGCQ that resulted from the analysis of the development sample study reported chapter 5.

Clinical Governance Climate Questionnaire (CGCQ)

We would like to know what it is like to work in your organisation

Thinking of the part of the organisation where you work, indicate whether you agree or disagree with each of the following statements. Please rate all the statements on the 5-point scale by circling the number that best represents your opinion.

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly

	<i>In the part of the organisation where I work...</i>	<i>(Circle)</i>
1	When things go wrong, there is an automatic assumption that 'someone is to blame'	1 2 3 4 5
2	Good practice ideas are shared with others outside the organisation	1 2 3 4 5
3	Error reporting systems are basically a stick to beat clinicians with	1 2 3 4 5
4	Critical appraisal skills training is available to those who want it	1 2 3 4 5
5	People involved in clinical incidents are made to feel guilty	1 2 3 4 5
6	Career development needs are addressed alongside strategic needs of the service	1 2 3 4 5
7	Staff appraisals are used to punish staff	1 2 3 4 5
8	People have a good knowledge of the skills of their colleagues	1 2 3 4 5
9	We have protected time for quality improvement activity	1 2 3 4 5
10	It is unsafe to be open and honest with colleagues	1 2 3 4 5
11	The emphasis is on how an incident happened, not who made the mistake	1 2 3 4 5
12	People who make mistakes are supported	1 2 3 4 5
13	We work together across teams to make quality improvements	1 2 3 4 5
14	We collect information on clinical risks	1 2 3 4 5
15	Technical help with evidence based practice is available	1 2 3 4 5
16	When there is an error, we look for failures in systems rather than blame individuals	1 2 3 4 5
17	Appraisal does not identify the real development needs of staff	1 2 3 4 5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly
In the part of the organisation where I work... (Circle)

18	Long-term planning for quality improvement gets lost in the day-to-day	1	2	3	4	5
19	Colleagues are dishonest with each other	1	2	3	4	5
20	Good practice stays isolated in pockets	1	2	3	4	5
21	Identified clinical risks simply remain unaddressed	1	2	3	4	5
22	Clinical risks are examined in a systematic way	1	2	3	4	5
23	People share practice issues with others in different parts of the organisation	1	2	3	4	5
24	There is no support to deliver service changes	1	2	3	4	5
25	There is no clear vision of what it is that the organisation is trying to achieve	1	2	3	4	5
26	We work in an atmosphere of blame	1	2	3	4	5
27	We don't collect information on the clinical risks that matter most	1	2	3	4	5
28	There is no training available in searching for research evidence	1	2	3	4	5
29	There are lots of quality improvement initiatives, but little real change	1	2	3	4	5
30	There is no common approach to risk management	1	2	3	4	5
31	Teams from different parts of the organisation share their good practice	1	2	3	4	5
32	People devote time to disseminating good practice	1	2	3	4	5
33	The first we know of quality improvements elsewhere in the organisation is when we feel their effects	1	2	3	4	5
34	Service improvements tend to be crisis-led	1	2	3	4	5
35	When a clinical risk is identified, there is always action to address it	1	2	3	4	5
36	Quality improvement is imposed from above rather than built from below	1	2	3	4	5
37	We systematically assess clinical risks	1	2	3	4	5
38	There is no time to get together to share ideas	1	2	3	4	5
39	People share a common vision of service delivery	1	2	3	4	5
40	There is mutual respect for everyone's contribution	1	2	3	4	5

1= Agree strongly, 2= Agree, 3= Neither, 4= Disagree, 5= Disagree strongly
In the part of the organisation where I work... (Circle)

41	There is pressure to 'solve' problems quickly rather than take time and do it properly	1	2	3	4	5
42	We don't address the accidents waiting to happen	1	2	3	4	5
43	Clinical risk policies are shared throughout the organisation	1	2	3	4	5
44	Clinical risk information is used routinely to inform decisions	1	2	3	4	5
45	People don't know what their colleagues expect of them	1	2	3	4	5
46	People don't seem to have shared service goals	1	2	3	4	5
47	There is time to reflect on practice	1	2	3	4	5
48	People don't know about good practice taking place in other parts of the organisation	1	2	3	4	5
49	Development needs are regularly assessed	1	2	3	4	5
50	Immediate pressures are always more important than quality improvement	1	2	3	4	5
51	Quality improvement activity is largely a response to external pressure	1	2	3	4	5
52	Risk assessment processes are updated in the light of clinical incidents	1	2	3	4	5
53	Colleagues don't seem to understand each others role's	1	2	3	4	5
54	When something fails, it is used as a learning opportunity	1	2	3	4	5
55	Everyone has the same standing, regardless of professional background	1	2	3	4	5
56	We react to problems, rather than try to prevent them	1	2	3	4	5
57	People are motivated to improve quality	1	2	3	4	5
58	There are few opportunities to use new skills learned as part of development	1	2	3	4	5
59	People are forced into making service changes, rather than encouraged to make them	1	2	3	4	5
60	People are highly motivated to make changes to clinical practice	1	2	3	4	5

Appendix 8: Skewness and kurtosis of CGCQ items pre- and post- transformation

This appendix lists skewness and kurtosis statistics for raw and transformed CGCQ indicator variables included in the factor analysis. As factor analysis results are optimised with normally distributed variables, those variables whose distributions could better approximate normal distribution following transformation were transformed, either by taking the square root of the variable or its base 10 logarithm.

1. Highly +vely skewed items: Newn= LG10 (Oldn)
2. +vely skewed items: Newn= SQRT (Oldn)
3. -vely skewed items: Newn = SQRT (6 - Oldn)
4. Highly -vely skewed items: Newn = LG10 (6 - Oldn)

'*' Denotes items with highly skewed distributions before transformation;

'+' Denotes items with poor discrimination excluded from analysis.

Item	S	K	+ve / -ve skewness	tS	tK	Transformation
1	-.202	-.983	-	-.142	-.854	3
2	-.550	-.576	-	.155	-.720	3
3	.305	-1.018	+	-.007	-.900	2
4	-.137	-1.127	-	.141	-.927	3
5	-.420	-.699	-	.029	-.821	3
6	-.683	-.132	-	.247	-.290	3
7	-.529	-.526	-	.177	-.502	3
+8*	-1.006	.478	--	-.056	.226	4
+9*	-1.049	1.259	--	-.206	.366	4
10	-.324	-.752	-	-.059	-.713	3
11	-.666	-.079	-	.201	-.384	3
12	-.992	.735	-	.476	.256	3
13	-.619	.236	-	.182	-.116	3
+14	-.779	.480	-	.254	-.232	3
15	-.107	-.958	-			None
16	-.666	-.405	-	.294	-.486	3
17	-.443	-.631	-	.084	-.727	3
+18	-.790	-.102	-	.353	-.465	3
+19	-.955	.470	-	.464	.121	3
20*	-1.081	1.230	--	.117	-.847	4
21	-.063	-.890	-			None
22	-.749	.223	-	.262	-.453	3
23	-.196	-.693	-			None
24	-.747	.367	-	.271	.151	3
25	-.427	-.424	-	-.006	-.487	3
26	-.328	-.741	-	-.044	-.633	3
27	-.593	-.064	-	.115	-.464	3
28	-.499	-.708	-	.142	-.697	3
29*	-1.108	1.447	--	.259	-1.082	4
30	-.502	-.606	-	.163	-.546	3
31	.281	-.830	+	-.095	-.823	2
32	-.653	-.011	-	.179	-.435	3
33	-.229	-.426	-	-.128	-.309	3
34	-.599	-.185	-	.146	-.374	3
35	-.413	.074	-	-.082	.008	3
+36	-.951	.693	-	.408	.012	3
+37*	-1.032	.484	--	.145	-.914	4

Item	S	K	+ve / -ve skewness	tS	tK	Transformation
38	-.379	-.319	-	.013	-.239	3
+39*	-1.206	2.223	--	-.086	-.303	4
40	-.825	.232	-	.336	-.157	3
41	-.737	.163	-	.244	-.272	3
42	-.909	.851	-	.414	.827	3
43	-.420	-.288	-	-.032	-.340	3
44	-.559	-.312	-	.133	-.370	3
+45*	-1.082	1.839	--	.092	-.917	4
46	-.438	.321	-	-.087	-.014	3
47	-.602	-.369	-	.208	-.324	3
48	.124	-.933	+			None
49	-.503	-.671	-	.126	-.631	3
50	-.665	.170	-	.163	-.238	3
51	-.820	.420	-	.337	.171	3
52	-.529	-.278	-	.116	-.276	3
53	-.619	-.037	-	.147	-.254	3
54	-.627	-.551	-	.328	-.604	3
55	-.603	-.259	-	.183	-.307	3
56	.197	-.916	+			None
57	-.545	-.461	-	.213	-.519	3
58	-.114	-.641				None
59	-.598	.371	-	.065	-.121	3
60	-.716	.515	-	.198	.214	3
+61*	-1.009	.529	--	.084	.645	4
62	-.321	-.258	-	.134	-.342	3
63	-.699	.399	-	.222	.184	3
64	-.229	-1.078	-	-.042	-.957	3
65	-.395	.773	-	.066	-.620	3
66	-.575	-.111	-	.107	-.335	3
67	-.423	-.203	-	-.037	-.242	3
68*	-.998	1.187	-	-.257	.638	4
69	-.268	-.948	-	-.035	-.784	3
70	-.414	-.887	-	.109	-.883	3
71	-.627	-.586	-	.336	-.594	3
+72*	-.934	1.728	--	-.287	-.229	4
73	-.472	-.549	-	.097	-.445	3
74	-.438	-.645	-	.098	-.497	3
75	-.661	-.015	-	.199	-.209	3
76	-.531	-.033	-	.075	-.98	3
77	-.472	-.607	-	.098	-.566	3
78	-.135	-.769				None
+79	-.804	.718	-	.241	.064	3
80	-.538	-.708	-	.172	-.791	3
81	-.770	-.043	-	.319	-.485	3
82	-.294	-.889	-	-.065	-.823	3
83	-.492	-.149	-	.023	-.278	3
+84	-.753	.772	-	.185	.340	3
+85*	-.981	1.883	--	.316	1.317	4
86	-.489	-.449	-	.075	-.472	3

Item	S	K	+ve / -ve skewness	tS	tK	Transformation
87	-.230	-.973	-	-.032	-.887	3
+88	-.706	-.262	-	.380	-.238	3
+89	-.579	-.719	-	.301	-.663	3
90	-.449	-.614	-	.055	-.639	3
91	-.332	-.891	-	.025	-.730	3
92	.055	-1.077	+			None
93	-.159	-.941	-			None
94	-.216	-.773	-	-.128	-.557	3
+95*	-1.010	2.241	--	-.308	-.042	4
96	-.370	-.936	-	.085	-.866	3
97	-.306	-.407	-	-.122	-.265	3
98	-.556	.066	-	.122	.028	3
99	-.196	-.878	-			None
+100*	1.064	1.759	++	-.158	-.281	1
101	-.403	-.507	-	.028	-.415	3
102	-.546	-.072	-	.169	-.199	3
103	-.098	-.646	-			None
104	-.248	-1.103	-	-.022	-.918	3
105	-.546	-.717	-	.255	-.717	3
106	-.498	-.596	-	.148	-.583	3
+107	-.930	.494	-	-.128	.501	4
108	-.493	-.542	-	.123	-.525	3
109	.068	1.099	+			None
110	-.318	-.828	-	-.012	-.715	3
+111*	-1.218	1.343	--	.123	1.427	4
112*	-1.001	.756	--	-.102	.697	4
+113*	-.970	.431	-	-.014	.758	4
+114	-.747	1.187	-	.129	.584	3
115	-.191	-.785	-			None
116	-.327	-.311	-	-.045	-.176	3
117	-.701	.466	-	.159	-.055	3
118	-.349	-.306	-	-.057	-.194	3
119	-.587	-.414	-	.316	-.595	3
+120*	-1.252	2.849	--	.128	-.834	4
121	-.476	-.569	-	.073	-.631	3
122	-.011	-1.015	-			None
123	-.604	-.368	-	.182	-.571	3
124	-.473	-.866	-	.196	-.748	3
125	-.295	-.927	-	.005	-.786	3
126	-.156	-1.096	-			None
127	-.974	1.416	-	.398	-.113	3
128	.007	-.998	+			None
129	-.469	-.600	-	.118	-.543	3
+130	-.891	.390	-	.481	.084	3
131	-.519	-.444	-	.153	-.392	3
132	-.375	-.581	-	.048	-.494	3
133	-.298	-.751	-	-.037	-.582	3
134	-.485	-.811	-	.196	-.632	3
135	-.443	-.535	-	.052	-.494	3

Item	S	K	+ve / -ve skewness	tS	tK	Transformation
136	-.855	-.097	-	.429	-.507	3
137	-.421	-.789	-	.078	-.735	3
138	.125	-.959	+			None
139	-.202	-.992	-	.084	-.850	3
140	-.374	-.817	-	.082	-.664	3
141	-.356	-1.089	-	.118	-.974	3
142	-.797	.833	-	.277	.519	3
143	-.826	.256	-	.368	.061	3
+144*	1.005	.865	--	-.168	.600	4
145	-.617	-.380	-	.243	-.365	3
146	-.411	-.828	-	.107	-.673	3
147	-.867	.267	-	.438	.065	3
148	-.057	-.670	-			None
149	-.704	.704	-	.135	.255	3
150	-.254	-.560	-	-.062	-.491	3
151	.324	-.776	+	-.048	-.685	2
152	-.108	1.184	-			None
153	-.568	-.277	-	.156	-.211	3
154	-.509	-.582	-	.166	-.486	3
155	-.361	-.790	-	.062	-.689	3
156	-.221	-.608	-	-.143	-.378	3
157	-.621	-.068	-	.212	-.171	3

Appendix 9: Complete rotated factor scores for six-factor PFA with Varimax rotation

This appendix gives complete details of the rotated six-factor matrix reported in Chapter 5, section 5.1.6.

Var	Factor 1	2	3	4	5	6	h ²
T87	.609	.220	.153	.144	.123	.164	.505
T154	.552	.174	.194	.261	.319	7.994E-02	.550
T128	-.552	-.274	-8.702E-02	-.152	-.156	-.330	.544
T138	-.547	-.153	-.158	-.172	-.153	-.255	.466
T125	.547	.237	.128	.282	.269	.200	.564
T139	.544	.234	.196	.223	.143	8.195E-02	.466
T94	.529	.290	.164	-6.651E-03	1.892E-02	.181	.424
T56	-.523	-.151	-.197	-.118	-.225	-.298	.488
T152	-.516	-.311	-.247	-.178	-.148	-.189	.513
T80	.510	.344	.174	.110	.226	.157	.496
T155	.495	.253	.295	.258	.227	.163	.541
T96	.490	.244	.361	.104	.190	.294	.563
T122	-.489	-2.978E-02	-.121	-.278	-2.898E-02	-.210	.371
T104	.486	7.265E-02	.169	3.508E-02	.316	.256	.437
T123	.481	.328	.331	.107	7.740E-02	5.118E-02	.480
T156	.475	.219	9.213E-02	.283	.253	.160	.452
T77	.468	.377	.235	8.262E-02	.298	.109	.524
T99	-.468	-.190	-.169	-.143	-.241	-.151	.385
T153	.463	.183	.173	.355	.255	.147	.491
T110	.458	.406	.272	.148	.115	.201	.524
T64	.442	.284	.178	.281	8.648E-02	.386	.544
T106	.434	.246	.105	.261	.131	.281	.425
T109	-.431	-.186	-.186	-7.937E-02	-.187	-.306	.390
T141	.429	.173	.152	.129	.315	.117	.372
T65	.414	.283	.178	.271	.178	.310	.484
T15	.408	.108	.311	.241	.102	.328	.455
T131	.408	.230	.100	.286	.306	8.028E-02	.404
T119	.407	3.220E-02	.191	5.886E-02	.261	7.509E-02	.261
T49	.404	.153	.160	.184	.238	-2.063E-03	.292
T82	.394	.130	.203	.136	-5.682E-02	.110	.251
T135	-.389	-.143	-.115	-.224	-8.377E-02	-.236	.309
T4	.386	7.481E-02	.130	.195	.128	.277	.297
T92	-.379	-.369	-.263	-.158	-4.094E-02	-.260	.457
T134	.372	.212	9.399E-02	.351	.369	.122	.449
T69	.370	.245	.293	.125	.133	.227	.360
T40	.359	.211	.247	.216	.153	.185	.339
T70	.355	.132	.277	.164	.320	.215	.407
T132	.350	-4.797E-02	.218	8.683E-02	.158	6.824E-02	.194
T76	.345	.187	2.753E-02	.216	.187	-4.881E-02	.220
T74	.314	7.787E-02	.278	.281	.216	.111	.317
T67	.104	.661	.259	9.846E-02	.156	.232	.603

Var	Factor 1	2	3	4	5	6	h ²
T50	3.614E-02	.636	.189	.112	.109	-1.193E-02	.466
T101	.203	.597	.229	8.367E-02	.112	.182	.503
T83	.347	.593	.345	.106	6.191E-02	7.428E-02	.612
T142	.175	.570	.210	.181	.131	.155	.474
T118	.281	.545	.177	.121	.129	.192	.475
T97	.180	.538	.325	.149	.116	.221	.512
T117	.152	.523	.188	9.061E-02	.175	-6.131E-02	.374
T90	.249	.499	.188	3.979E-02	.122	3.996E-02	.364
T150	.123	.460	-4.941E-02	7.231E-02	.309	.296	.421
T147	.262	.452	.291	.249	.276	.175	.526
T66	.340	.439	.345	.116	.108	.162	.479
T43	.297	.414	.269	.351	8.966E-02	3.195E-02	.430
T102	.195	.410	-1.938E-03	.232	.247	.310	.408
T137	.343	.390	4.707E-02	-6.450E-03	.171	.212	.356
T103	-.232	-.390	-6.955E-02	-.165	-.199	-.311	.381
T115	-.384	-.385	-.164	-.187	-.175	-.374	.517
T68	.196	.369	.197	.360	.303	.155	.483
T98	.224	.366	.347	.289	.193	.102	.422
T33	6.327E-02	.348	.247	.249	3.446E-02	.161	.259
T148	-.176	-.340	-9.875E-02	-7.731E-02	-.252	-.322	.335
T63	.233	.337	7.429E-02	.305	.285	.295	.458
T13	3.828E-02	.334	.333	.109	.190	.259	.366
T112	.287	.316	.264	.297	.289	.276	.391
T27	.257	.280	.646	.106	.122	.203	.629
T34	9.335E-02	.359	.616	.148	.187	.172	.604
T81	.361	.250	.602	.109	.146	.180	.621
T1	.236	.105	.601	7.064E-02	.137	.269	.524
T22	.196	.348	.544	.126	7.054E-02	1.680E-02	.476
T29	8.448E-02	.255	.535	8.634E-02	.315	.131	.482
T53	.138	.400	.520	.209	.164	.218	.567
T32	.256	9.892E-02	.493	.262	.174	.147	.439
T42	.161	.270	.481	.154	.128	.124	.386
T127	.256	.175	.426	5.793E-02	.334	-.106	.458
T11	.185	.291	.418	-3.010E-02	.106	.177	.341
T51	.242	.332	.411	.317	.197	.263	.553
T62	.140	4.868E-02	.393	7.566E-02	4.777E-02	1.013E-02	.181
T143	.317	.284	.390	.316	.300	8.973E-02	.514
T16	.340	.148	.389	.300	.112	.246	.446
T60	.103	.273	.388	.346	.196	6.676E-02	.367
T121	.338	.197	.372	.106	.251	4.441E-02	.363
T6	7.762E-02	.358	.367	.218	.109	.189	.396
T35	.206	.176	.334	.300	-7.501E-02	7.319E-02	.309
T54	.323	9.525E-02	5.786E-02	.603	.175	.156	.494
T30	8.589E-02	8.623E-02	2.514E-02	.562	.190	.136	.386
T146	.451	8.767E-02	8.552E-02	.538	.245	.182	.601
T124	.458	.123	.141	.471	.257	5.821E-02	.535
T151	-.343	-8.737E-02	-.130	-.459	-.154	-.241	.435
T108	.354	.161	.298	.427	.130	.313	.538
T59	.201	1.354E-02	.376	.424	-1.714E-03	9.211E-02	.370
T38	3.152E-02	.187	.287	.424	3.136E-02	.148	.302
T75	.144	.276	.312	.394	.212	6.785E-02	.390
T105	.289	6.875E-02	.250	.387	.167	.300	.423
T47	.237	9.711E-02	.245	.384	.239	.348	.451

Var	Factor 1	2	3	4	5	6	h^2
T48	-.290	-8.353E-02	-.273	-.372	-1.657E-03	-.237	.361
T41	.255	.343	.250	.371	.211	2.362E-02	.382
T157	.170	.272	.305	.357	.222	2.691E-02	.336
T23	-.200	-.174	-.137	-.344	-4.800E-02	-.259	.267
T24	4.762E-02	.196	6.807E-02	.329	.310	.194	.298
T86	.244	.258	.144	4.039E-02	.521	1.280E-02	.420
T25	-3.532E-02	8.687E-02	5.941E-02	.133	.512	.176	.324
T52	.162	.323	.196	5.809E-02	.503	.232	.480
T28	.195	.233	.193	.154	.458	.247	.424
T129	.293	.246	.150	.133	.441	.267	.452
T31	-.148	-2.159E-02	-.157	-5.228E-02	-.429	-.404	.397
T55	.278	.129	.278	.229	.421	.105	.412
T126	-.299	-1.288E-02	-3.010E-02	5.100E-02	-.420	-.216	.316
T26	.265	.189	.282	.216	.420	.268	.475
T57	.269	9.841E-02	.300	.210	.419	.172	.422
T145	.297	.127	.120	.366	.416	.180	.455
T116	.193	.308	2.850E-02	.122	.413	.263	.377
T149	.254	.206	.144	.222	.390	-8.325E-02	.354
T46	7.780E-02	.191	.330	.169	.383	.165	.343
T133	.298	.269	.145	.250	.346	.271	.434
T17	.277	.334	.204	2.271E-02	.343	.252	.417
T136	.130	5.710E-02	7.273E-02	.136	.303	-3.258E-02	.171
T140	.257	.228	9.930E-02	.175	.280	.186	.272
T71	.287	.254	.147	.267	.193	.522	.549
T10	.220	.193	.155	.143	.232	.476	.410
T93	-.375	-.310	-3.885E-02	-.194	-.370	-.473	.637
T3	-.349	-6.936E-02	-.195	-.158	-.119	-.470	.409
T91	.359	.362	.158	.262	.182	.467	.604
T44	.317	.207	.196	.350	.184	.442	.533
T73	.295	.347	.220	.227	.114	.407	.473
T21	-.248	-7.393E-02	-.180	-5.210E-02	-.349	-.397	.386
T12	.220	.256	.228	.161	.150	.391	.362
T2	.349	.272	.239	.167	.153	.374	.435
T7	7.740E-02	.115	.231	.289	.149	.361	.314
T58	-.279	-.272	-2.442E-02	-.195	-.177	-.357	.376
T78	-.135	-4.598E-02	-.110	-.212	7.149E-03	-.260	.142
% variance	10.6	8.1	7.0	5.8	5.7	5.5	
% covariance	24	19	16	14	14	13	

Appendix 10: Instructions to Q-sort participants

The following pages detail the instructions given to each Q-sort participant. They include details of the overall aim of the exercise, along with step-by-step instructions for completion of the task. Participants received a 'participant pack', consisting of two sealed envelopes containing the items on cards, and two scorecards for recording the ID numbers of the cards allocated under each of the six marker headings for each sort. Participants also had two adjacent tables with marker headings from '1' to '6' on each, on which to conduct the sorts.

The Q-sort process

The object of the Q-sort is to sort the items into *six piles* representing increasingly difficult achievements in clinical governance. At the end of the procedure, the items under heading '1' will be those considered to represent the easiest achievements, representing basic competence in clinical governance. The items under successive headings are judged to be increasingly difficult, through to those under heading '6' which are judged to be the most difficult to achieve, representing a high level of competence in clinical governance. **There are no right answers** - the purpose of the procedure is to gather opinions on the possible ordering of the statements.

Instructions to participants

1. Open 'pack 1' and remove all of the cards. Read the cards once in order to become familiar with them, ignoring the identification number, which is for administrative purposes.
2. Read through the cards a second time, and select the 12 that you think are the most difficult to achieve, representing a high level of competence in clinical governance, and place them vertically, under the '6' marker. The order of the cards under the marker is not important: all 12 items will receive the same score.
3. Read the remaining cards, and select the 12 that you think are the easiest to achieve, representing the most basic competency in clinical governance, and place them vertically, under the '1' marker. Again, the specific order does not matter.
4. Next, pick the 12 cards from those remaining that represent the most difficult to achieve, but which are not as difficult as the 12 already placed under the '6' marker, and place them under the '5' marker. On second thought, you might decide that a card selected for '5' is more difficult than one under '6'. If so, feel free to switch it with another at this or any other time.
5. Repeat the above process for the headings '2', '4' and finally '3', placing 12 cards under each heading and reclassifying individual items as you feel necessary until you are happy with the items under each heading.
6. Once you are happy, repeat the above process on the cards contained in 'pack 2'.
7. After completing the process for pack 2, look over the two allocations. Compare the cards across the two allocations. If for example you feel that one of the cards under the heading '5' in one allocation is more difficult than one under '6' in the other, feel free to swap them.
8. Keep comparing and swapping the cards from the two allocations until you are happy with the result.
9. Finally, record the identification number of each card under the appropriate heading on the Q-sort scorecard, available from the moderator.

Thank you very much for your help and co-operation, which is greatly appreciated

First Q-sort scorecard

Please record the ID number of the 12 cards under each of the six headings for
the first sort

1	2	3	4	5	6

Name:

Second Q-sort scorecard

Please record the ID number of the 12 cards under each of the six headings for
the second sort

1	2	3	4	5	6

Name:

Appendix 11: Q-sort results

The tables below list the number of times participants rated each item in each category from '1' (low) to '6' (high). Participants had six options in order to facilitate ratings between items; for analytical purposes, ratings of 1 or 2 were considered 'low', 3 or 4 were considered 'medium', and 5 or six 'high'. Column 2 identifies those items retained for the final schedule.

Old No.	In?	New No.	Statement	Scores:	Low						High	Level
					1	2	3	4	5	6		
			Task 1: Benchmarking practice									
1.1			People are encouraged to be innovative in their approach to clinical practice		1	1	1	1	1	2		
1.2			Quality improvement activity is supported		3	1	2	1	0	0		
1.3			Once agreement on the need for change has been reached, practice changes are made		2	0	1	1	1	2		
1.4			Staff are encouraged to identify problems and solutions		2	2	1	0	1	1		
1.5			Quality improvement activity is backed by resources [time and money]		1	2	1	0	1	2		
1.6			Quality problems are addressed with structured problem-solving techniques		1	2	1	2	0	1		
1.7			Staff compare their care processes with those of other providers		0	1	2	1	2	1		
1.8	☒	2	Staff identify best practice through benchmarking against other providers		0	0	3	1	2	1	Med	
1.9	☒	3	Where appropriate, staff modify their care processes to reflect the best practice of others		1	1	1	3	1	0	Med	
1.10			Staff evaluate the impact of any changes they make to practice following a benchmarking exercise		0	0	0	3	2	2		

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
Task 2: Using research evidence							
2.1			There is good access to research evidence via paper and/or electronic means		2	0 3 0 1 1	
2.2			Research evidence is critically appraised before use, not just accepted at face value		0	2 2 2 1 0	
2.3			Research evidence is used in quality improvement activity		0	2 2 2 1 0	
2.4	☒	4	Research evidence is regularly reviewed and discussed		1	1 2 3 0 0	Med
2.5			Changes in practice draw on available research		0	1 1 3 1 1	
2.6	☒	5	Training in evidence based practice and critical appraisal is available to staff		7	0 0 0 0 0	Low
2.7			There are agreed processes for consideration of information and decision-making on actions		0	1 2 1 2 1	
2.8	☒	6	Research Information is used consistently to inform our approach to quality improvement		0	0 3 3 1 0	Med
Task 3: Discussing clinical issues							
3.1			Clinical colleagues share expectations of a standard of clinical care		1	0 2 1 1 2	
3.2	☒	8	Discussions on clinical issues are not dominated by any single profession		0	0 2 1 2 2	High
3.3			There are systems and processes in place to encourage discussion of clinical issues		1	2 2 1 0 1	
3.4			There is an open and honest atmosphere		0	2 0 1 0 4	
Task 4: Auditing clinical practice							
4.1			Data from complaints, adverse events and risk management systems is used to inform the audit agenda		0	1 2 1 1 2	
4.2	☒	9	Staff participate in audit activity		4	3 0 0 0 0	Low

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
4.3			Audit methods are well understood		1 4 1 1 0 0		
4.4			Clinical audit is change focussed- it is used to improve clinical quality		1 0 3 1 1 1		
4.5			People participate in national audits such as confidential enquiries		4 0 2 0 1 1		
4.6			Audits result in documented practice change		1 0 0 4 2 0		
4.7	☒	10	Training in audit is available		7 0 0 0 0 0		Low
4.8	☒	11	Topics for audit are selected according to their potential impact on care quality		2 3 1 1 0 0		Low
Task 5: Using performance indicators							
5. 1			Performance indicators are selected following widespread debate within teams		0 1 2 0 2 2		
5. 2	☒	24	There is good access to agreed clinical performance indicators		0 2 4 0 1 0		Med
5.3	☒	25	Clinical indicators are used to reflect on, review and integrate services		1 0 0 4 1 1		Med
5. 4			Time is made to examine clinical performance indicators		1 1 1 2 2 0		
5. 5			There is a capacity to analyse and interpret clinical performance indicator data		0 2 3 0 2 0		
5. 6			Clinical performance indicators are successfully used to inform quality improvement initiatives		0 1 2 1 1 2		
Task 6: Using risk management systems							
6. 1			There is a robust risk management system		2 1 0 0 0 4		
6. 2	☒	12	Risk management data is regularly reviewed and discussed		2 2 0 1 1 1		Low
6. 3			Staff clearly understand the risk management process		1 1 2 1 1 1		

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
6. 4	☒	13	Following identification of a problem from risk data, clinical quality is improved		1 0 0 2 2 2		High
6. 5			The training implications of specific clinical risks are noted, and training provided		0 2 1 1 2 1		
6. 6	☒	14	Clear action plans are developed in response to identified clinical risks		0 1 2 2 1 1		Med
6. 7			Following implementation, action plans to address clinical risk are reviewed		1 1 1 0 2 2		
6. 8	☒	15	Staff receive training in using the risk management system		5 1 0 0 0 1		Low
Task 7: Using adverse event systems							
7. 1			There is a robust adverse event reporting system		2 1 0 0 1 3		
7. 2	☒	16	Data from the adverse event reporting system is regularly reviewed and discussed		1 3 0 1 2 0		Low
7. 3			Following implementation, action plans to address adverse events are reviewed		0 3 0 0 2 2		
7. 4			Training opportunities are provided if shown to be necessary by investigations into adverse events		2 2 1 1 1		
7. 5			Adverse events inform the organisations training and development plan		0 0 0 3 1 3		
7. 6	☒	17	Clear action plans are developed in response to adverse incidents		0 2 1 3 0 1		Med
7. 7	☒	18	Following identification of a problem from adverse events data, clinical quality is improved		1 0 1 1 2 2		High
7. 8	☒	19	Staff are trained to use adverse event systems		5 0 0 1 1 0		Low
Task 8: Using complaints systems							
8. 1			There is a robust complaints reporting system		2 1 0 0 2 2		
8. 2			Data from the complaints reporting system is regularly reviewed and discussed		3 1 2 0 1 0		

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
8. 3	☒	20	Complaints are collated		5	1	Low
8. 4			There are clear links between complaints monitoring and training and opportunities		0	1	
8. 5			Following discussion of complaints data, action plans are prepared and implemented		0	2	
8. 6			Following implementation, action plans to address complaints are reviewed		1	2	
8. 7	☒	21	Following identification of a problem from complaints data, clinical quality is improved		1	0	High
Task 9: Learning from complaints, adverse events and near misses							
9. 1			Complaints / adverse events / near misses are used as an opportunity to improve practice		0	0	
9. 2	☒	22	There is a 'no-blame' culture around reporting adverse events and near misses		0	0	High
9. 3			Following adverse events, there is a detailed analysis of the wider context of contributing factors		0	0	
9. 4			Action plans are drawn up in the light of a detailed analysis of the background to adverse events		0	0	
9. 5			Changes in practice resulting from action plans to address adverse events are evaluated		0	2	
Task 10: Raising issues in confidence							
10.1	☒	7	Clinical issues are raised for discussion		4	2	Low
10. 2	☒	23	Staff can raise clinical concerns about their colleagues in confidence		1	1	High
Task 11: Professional development for staff							
11. 1	☒	26	Staff have development plans which identify training and development opportunities		2	2	Low

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
11.2	☒	27	Training identified in staff development plans matches individual needs to organisational needs	1 2 2 1 1 0			Med
11.3			Training identified in staff development plans is delivered	2 1 1 2 1 0			
11.4			There are opportunities for mentoring and / or clinical supervision	0 3 3 1 0 0			
11.5	☒	28	New skills gained through development activity are used in clinical settings	0 0 2 1 3 1			High
11.6			There are shared standards of acceptable clinical practice	1 1 1 0 2 2			
11.7			In managing poor performance, the emphasis is on prevention and support through education and development opportunities	0 2 3 0 1 1			
11.8			Under-performance in clinical staff is addressed	0 2 1 2 2 0			
11.9			There are local arrangements for detecting and addressing performance problems before they escalate	0 1 2 1 2 1			
11.10			Serious clinical under-performance is swiftly referred to appropriate external bodies	1 2 1 1 0 2			
Task 12: Appraising staff							
12.1	☒	29	There is an annual staff appraisal process for most staff	4 2 0 1 0 0			Low
12.2	☒	30	An agreed work and development programme is used as the basis of staff appraisal	0 0 3 2 2 0			Med
12.3			Staff appraisals occur at least annually	3 1 1 2 0 0			
12.4	☒	31	Staff appraisal is used as an opportunity to reflect on progress and plan future development	0 2 2 3 0 0			Med
12.5			360° feedback is used in staff appraisals- people get feedback from both subordinates and their managers	0 0 1 2 3 1			
12.6			The time and resources required for the staff appraisal process are acknowledged and made available	2 0 1 1 1 2			

Old No.	In?	New No.	Statement	Scores:						Low	High	Level
Task 13: Accountability for clinical quality												
13.1	☒	32	There is an executive director with responsibility for developing the clinical governance agenda within the trust	7	0	0	0	0	0	0	0	Low
13.2			Clinical areas collate information on their clinical governance activity	2	2	1	1	0	1			
13.3	☒	33	Quality improvement activity is planned and is included in a service development plan	2	1	1	3	0	0			Med
13.4			Staff job descriptions detail staff responsibilities for clinical governance	3	2	1	1	0	0			
13.5	☒	34	Clinical areas have a nominated lead for clinical governance	7	0	0	0	0	0	0	0	Low
Task 14: Organising CG committee structures												
14.1	☒	35	There is a formal clinical governance committee, reporting to the board	7	0	0	0	0	0	0	0	Low
14.2			The clinical governance committee requests and examines information on quality improvement, risk management and clinical performance activities across the trust	3	1	0	2	1	0			
14.3			The clinical governance committee recommends actions based on information requested and received	3	2	1	0	1	0			
14.4	☒	36	There are local arrangements to collate information for the clinical governance committee	5	0	0	0	1	1			Low
14.5			Actions recommended by the clinical governance committee take place	0	3	1	1	1	1			
14.6	☒	37	Organisation-wide clinical governance systems are underpinned by systems in clinical areas	2	0	2	2	0	1			Med
Task 15: Linking development activity to the clinical governance agenda												
15.1			The national clinical governance agenda is known and well understood by staff	0	0	3	2	1	1			

Old No.	In?	New No.	Statement	Scores:	Low	High	Level				
15.2	☒	38	Decisions about service developments or cutbacks are made on clear criteria		0	0	3	2	1	1	Med
15.3			Service development activities are reported to the appropriate clinical governance committees in accordance with agreed procedures		2	2	3	0	0	0	
15.4			New service developments are evaluated explicitly as part of their implementation		0	0	2	2	2	1	
15.5			Planning for quality improvement is incorporated into trust development plans		2	2	2	1	0	0	
15.6	☒	39	Local and national priorities from NSFs and HImPs are used to priorities service development		0	5	1	1	0	0	Low
15.7			There is a programme of service development activity		2	1	2	1	1	0	
15.8			The case for new procedures or services is made with reference to research evidence		0	0	2	1	3	1	
Task 16: Using NSFs and guidelines											
16.1			Dissemination strategies are designed around specific national service frameworks		0	3	1	1	1	1	
16.2			NSF implementation strategies are developed and delivered		1	1	2	1	2	0	
16.3			There is good access to electronic and paper sources of relevant information on NSFs		3	1	2	1	0	0	
16.4			There are agreed monitoring and review processes for implementation of NSFs		0	4	0	2	1	0	
16.5	☒	40	NSF implementation is integrated with business planning and quality improvement programmes		2	3	0	2	0	0	Low
16.6			Guidelines / protocols are reviewed at agreed time intervals		2	2	2	1	0	0	
16.7			There are clear procedures for disseminating externally generated guidelines		2	2	2	0	1	0	
16.8			Staff have good access to electronic and paper practice guidelines		2	1	2	2	0	0	
16.9			There are agreed processes for consideration of externally generated guidelines		2	2	0	2	1	0	

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
16.10			Local adaptation of guidelines follows an agreed formula		0 2 3 1 1 0		
16.11			Once accepted, the adherence to guidelines is evaluated		0 1 2 1 3 0		
16.12	☒	41	External guidelines are critically appraised before local adoption or development		0 0 4 1 1 1	Med	
16.13	☒	42	Clinical protocols are shared with staff who work outside this organisation		0 0 0 1 2 4	High	
Task 17: Leadership and teambuilding							
17.1	☒	43	The organisation shares a common vision for clinical governance		1 0 1 1 0 4	High	
17.2			There are agreed clear service objectives		3 0 1 1 0 2		
17.3			There are clear organisational objectives and priorities		3 1 0 0 0 3		
17.4			There is a clear link between local service aims and those of the wider organisation		1 1 1 2 2 0		
17.5	☒	44	Leadership skills are identified and developed through leadership training programmes		1 2 0 4 0 0	Med	
17.6			The skills of colleagues are known and recognised		0 1 2 1 2 1		
17.7	☒	45	Clinical teams receive performance feedback		0 1 1 4 1 0	Med	
17.8	☒	46	Clinical teams respond to changes in their environment by reorganising their work processes		0 1 0 0 5 1	High	
17.9			Staff engage in robust discussions without rancour		0 1 1 0 3 2		
17.10			Staff have clear roles and responsibilities		1 2 1 3 0 1		
17.11	☒	47	Staff have clear and shared objectives		1 1 0 0 3 2	High	
17.12							

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
Task 18: Involving service users in service development							
18.1	☒	48	There are clear processes for the involvement of service users in service development		0	1 0 1 1 4	High
18.2			Induction and development training is provided for users involved in service development		0	1 3 1 1 1	
18.3	☒	49	There are clear criteria for identifying and establishing user involvement groups		0	1 0 1 2 3	High
18.4			Service evaluation activity routinely includes service user perspectives		0	0 2 1 1 3	
18.5			Service users are involved in the review of data from the risk management, complaints and adverse events systems		0	1 1 0 5 0	
18.6	☒	50	Service improvement activity focuses on the patient experience of care		0	0 0 1 2 4	High
Task 19: Collaborating with local health and social care partners							
19.1			Staff have a knowledge of the local health and social care system relating to their service areas		0	2 1 1 1 2	
19.2			There is regular and active participation in joint fora for planning and development in clinical governance		0	1 1 1 2 2	
19.3			There is respect for the contributions of all stakeholders in collaborations with local health and social care agencies		0	0 0 0 3 4	
19.4			There is clarity of roles between local health and social care agencies		0	0 0 1 1 5	
19.5	☒	51	Local health and social care agencies work jointly on clinical governance issues		0	0 1 0 1 5	High
19.6			Key local stakeholders in health and social care agencies are identified, and good working relationships have been formed		0	0 1 0 3 3	High
19.7	☒	52	Local partnerships with health and social care agencies have clear, shared purposes		0	0 0 1 0 6	

Old No.	In?	New No.	Statement	Scores:	Low	High	Level
			Task 20: Developing organisational learning				
20.1			Staff critically reflect on their practice and learn from experiences		0 1 0 3 2 1		
20.2	☒	1	Teams work together across boundaries to address quality issues		0 1 0 2 1 3		High
20.3			Changes in practice resulting from quality improvement activity are shared with colleagues from across the organisation		0 1 2 1 2 1		
20.4			Staff conduct audits of their current practice against standards of best practice, to check that services comply with standards		1 3 2 1 0 0		
20.5	☒	53	Staff question what they are doing, and are able to develop new and innovative models of service		0 1 0 0 3 3		High
20.6	☒	54	Staff evaluate the best ways of training and learning from experience		0 0 0 0 4 3		High



**THE UNIVERSITY
OF BIRMINGHAM**

Organisational Progress in Clinical Governance (OPCG)

**A variance measure of organisational progress
in clinical governance**

© Tim Freeman HSMC 2001

Organisational Progress in Clinical Governance (OPCG)

A) Improving quality	competency item(s)
Benchmarking practice	2-3
Using research evidence	4-6
Discussing clinical issues	8
Auditing clinical practice	9-11
Using performance indicators	24-25
 B) Managing risks	
Using risk management systems	12-15
Using adverse events systems	16-19
Using complaints systems	20-21
Learning from complaints, adverse events and near misses	22
Raising issues in confidence	7, 23
 C) Improving performance	
Professional development for staff	26-28
Appraising staff	29-31
 D) Corporate accountability	
Accountability for clinical quality	32, 34
Organising clinical governance committee structures	35-37
 E) Leadership and collaboration	
Linking development activity to the clinical governance agenda	33, 38-39
Using National Service Frameworks (NSFs) & guidelines	40-43
Leadership and team building	44-47
Involving service users in service development	48-50
Collaborating with local health and social care partners	51-52
Developing organisational learning	1, 53-54

The aim of this questionnaire is to obtain your views about progress in clinical governance in your organisation.

The OPCG schedule is designed to canvass opinion on progress in clinical governance across a number of domains, including quality improvement processes, corporate accountability and collaboration with partner agencies. We seek your views on the relative importance of the elements of clinical governance, as well as your assessment of your Trust's success in integrating those elements into its working practices.

Some additional points

5. Funded by the NAO, the study is part of a linked series of research activities into clinical governance. Findings will be reported to the Public Accounts Committee.
6. We give an assurance that individuals responding to this questionnaire will not be identifiable in any reports of the findings. Your responses will be held in complete confidence by the research team at the Health Services Management Centre, University of Birmingham.
7. The statements contained in the questionnaire do not necessarily represent the views of HSMC.
8. **There are no right or wrong answers. We are simply interested in your views.**

Thank you for giving us some of your valuable time.

Initial questions

CODE

We need to know whether your organisation has undergone a visit from the Commission for Health Improvement (CHI), and whether you are a member of the trust board

1. Has your trust been subject to a CHI review?

Yes ☐1

No ☐2

2. Are you a member of the trust board?

Yes ☐1

No ☐2

The OPGC items

The remainder of the schedule asks for your opinion of your organisation's achievement against a series of statements reflecting good practice in clinical governance. For each statement, we would like you to indicate the current level of **achievement** of the organisation. We would also like you to indicate the level of **importance** you attach each statement. For example, the first statement is 'teams work across boundaries to address quality issues'. If you think that that this is not currently happening in your organisation, and that it is moderately important for clinical governance that it does happen, you would circle '1' under 'Achievement', and '3' under 'Importance'.

Organisational Progress in Clinical Governance (OPCG)

	Achievement						Importance					
	Low			High			Low			High		
1 Teams work across boundaries to address quality issues	1	2	3	4	5	6	1	2	3	4	5	6
2 Staff identify best practice through benchmarking against other providers	1	2	3	4	5	6	1	2	3	4	5	6
3 Where appropriate, staff modify their care processes to reflect the best practice of others	1	2	3	4	5	6	1	2	3	4	5	6
4 Research evidence is regularly reviewed and discussed	1	2	3	4	5	6	1	2	3	4	5	6
5 Training in evidence-based practice and critical appraisal is available to staff	1	2	3	4	5	6	1	2	3	4	5	6
6 Research information is used consistently to inform our approach to quality improvement	1	2	3	4	5	6	1	2	3	4	5	6
7 Clinical issues are raised for discussion	1	2	3	4	5	6	1	2	3	4	5	6
8 Discussions on clinical issues are not dominated by any single profession	1	2	3	4	5	6	1	2	3	4	5	6
9 Staff participate in clinical audit activity	1	2	3	4	5	6	1	2	3	4	5	6
10 Training in clinical audit is available to staff	1	2	3	4	5	6	1	2	3	4	5	6
11 Topics for audit are selected according to their potential impact on care quality	1	2	3	4	5	6	1	2	3	4	5	6
12 Risk management data is regularly reviewed and discussed	1	2	3	4	5	6	1	2	3	4	5	6
13 Following identification of a problem from risk data, clinical quality is improved	1	2	3	4	5	6	1	2	3	4	5	6
14 Clear action plans are developed in response to identified clinical risks	1	2	3	4	5	6	1	2	3	4	5	6
15 Staff are trained to use the risk management system	1	2	3	4	5	6	1	2	3	4	5	6
16 Adverse incident data is regularly reviewed and discussed	1	2	3	4	5	6	1	2	3	4	5	6
17 Clear action plans are developed in response to adverse incidents	1	2	3	4	5	6	1	2	3	4	5	6
18 Following identification of a problem from adverse events data, clinical quality is improved	1	2	3	4	5	6	1	2	3	4	5	6

	Achievement						Importance					
	Low			High			Low			High		
19 Staff are trained to use adverse event systems	1	2	3	4	5	6	1	2	3	4	5	6
20 Complaints are collated	1	2	3	4	5	6	1	2	3	4	5	6
21 Following identification of a problem from complaints data, clinical quality is improved	1	2	3	4	5	6	1	2	3	4	5	6
22 There is a 'no-blame' culture around reporting adverse events and near misses	1	2	3	4	5	6	1	2	3	4	5	6
23 Staff can raise clinical concerns about their colleagues in confidence	1	2	3	4	5	6	1	2	3	4	5	6
24 There is good access to agreed clinical performance indicators	1	2	3	4	5	6	1	2	3	4	5	6
25 Clinical indicators are used to reflect on, review and integrate services	1	2	3	4	5	6	1	2	3	4	5	6
26 Staff have development plans which identify training and development opportunities	1	2	3	4	5	6	1	2	3	4	5	6
27 Training identified in staff development plans matches individual needs to organisational needs	1	2	3	4	5	6	1	2	3	4	5	6
28 New skills gained through development activity are used in clinical settings	1	2	3	4	5	6	1	2	3	4	5	6
29 There is an annual staff appraisal process for most staff	1	2	3	4	5	6	1	2	3	4	5	6
30 An agreed work and development programme is used as the basis of staff appraisal	1	2	3	4	5	6	1	2	3	4	5	6
31 Staff appraisal is used as an opportunity to reflect on progress and plan future development	1	2	3	4	5	6	1	2	3	4	5	6
32 There is an executive director with responsibility for developing the clinical governance agenda	1	2	3	4	5	6	1	2	3	4	5	6
33 Service delivery plans include quality improvement activity	1	2	3	4	5	6	1	2	3	4	5	6
34 Clinical areas have a nominated clinical governance lead	1	2	3	4	5	6	1	2	3	4	5	6
35 There is a formal clinical governance committee, reporting to the board	1	2	3	4	5	6	1	2	3	4	5	6
36 There are local arrangements to collate information for the clinical governance committee	1	2	3	4	5	6	1	2	3	4	5	6
37 Organisation-wide clinical governance systems are underpinned by systems in clinical areas	1	2	3	4	5	6	1	2	3	4	5	6

	Achievement						Importance					
	Low			High			Low			High		
38 Decisions about service developments or cutbacks are made on clear criteria	1	2	3	4	5	6	1	2	3	4	5	6
39 Local and national priorities from NSFs and HImPs are used to priorities service development	1	2	3	4	5	6	1	2	3	4	5	6
40 NSF implementation is integrated with business planning and quality improvement programmes	1	2	3	4	5	6	1	2	3	4	5	6
41 External guidelines are critically appraised before local adoption or development	1	2	3	4	5	6	1	2	3	4	5	6
42 Clinical protocols are shared with staff who work outside this organisation	1	2	3	4	5	6	1	2	3	4	5	6
43 The organisation shares a common vision for clinical governance	1	2	3	4	5	6	1	2	3	4	5	6
44 Leadership skills are identified and developed through leadership training programmes	1	2	3	4	5	6	1	2	3	4	5	6
45 Clinical teams receive performance feedback	1	2	3	4	5	6	1	2	3	4	5	6
46 Clinical teams respond to changes in their environment by reorganising their work processes	1	2	3	4	5	6	1	2	3	4	5	6
47 Staff have clear and shared objectives	1	2	3	4	5	6	1	2	3	4	5	6
48 There are clear processes for involving service users in service development	1	2	3	4	5	6	1	2	3	4	5	6
49 There are clear criteria for establishing user involvement groups	1	2	3	4	5	6	1	2	3	4	5	6
50 Service improvement activity focuses on the patient experience of care	1	2	3	4	5	6	1	2	3	4	5	6
51 Local health and social care agencies work jointly on clinical governance issues	1	2	3	4	5	6	1	2	3	4	5	6
52 Local partnerships with health and social care agencies have clear, shared purposes	1	2	3	4	5	6	1	2	3	4	5	6
53 Staff question what they are doing, and are able to develop new and innovative models of service	1	2	3	4	5	6	1	2	3	4	5	6
54 Staff evaluate the best ways of training and learning from experience	1	2	3	4	5	6	1	2	3	4	5	6

Thank you for completing this questionnaire

We would be most grateful if you would return it to us here at HSMC in the envelope provided using the pre-paid slip **within two weeks of receipt of this form**. The return address, which is printed on the pre-paid slip enclosed, is repeated here:

Ms. Wendy Spurr
Health Services Management Centre
University of Birmingham
Park House
40 Edgbaston Park Road
Birmingham B15 2RT

If you would like further information on the study, then please do not hesitate to contact me by e-mail, telephone or letter:

Mr. T. Freeman

Lecturer

Tel: [REDACTED]

e-mail: [REDACTED]

Address: as above

Once again, many thanks for your time and co-operation; it is very much appreciated.

Appendix 13: OPGC items by structure, process or outcome

'Structure' items

- 5 Training in evidence-based practice and critical appraisal is available to staff
- 9 Staff participate in clinical audit activity
- 10 Training in clinical audit is available to staff
- 20 Complaints are collated
- 24 There is good access to agreed clinical performance indicators
- 29 There is an annual staff appraisal process for most staff
- 32 There is an executive director with responsibility for developing the clinical governance agenda
- 34 Clinical areas have a nominated clinical governance lead
- 35 There is a formal clinical governance committee, reporting to the board
- 36 There are local arrangements to collate information for the clinical governance committee
- 37 Organisation-wide clinical governance systems are underpinned by systems in clinical areas

'Process' items

- 1 Teams work across boundaries to address quality issues
- 2 Staff identify best practice through benchmarking against other providers
- 4 Research evidence is regularly reviewed and discussed
- 7 Clinical issues are raised for discussion
- 8 Discussions on clinical issues are not dominated by any single profession
- 11 Topics for audit are selected according to their potential impact on care quality
- 12 Risk management data is regularly reviewed and discussed
- 14 Clear action plans are developed in response to identified clinical risks
- 15 Staff are trained to use the risk management system
- 16 Adverse incident data is regularly reviewed and discussed
- 17 Clear action plans are developed in response to adverse incidents
- 19 Staff are trained to use adverse event systems
- 22 There is a 'no-blame' culture around reporting adverse events and near misses
- 23 Staff can raise clinical concerns about their colleagues in confidence
- 26 Staff have development plans which identify training and development opportunities
- 27 Training identified in staff development plans matches individual needs to organisational needs
- 30 An agreed work and development programme is used as the basis of staff appraisal
- 31 Staff appraisal is used as an opportunity to reflect on progress and plan future development
- 33 Service delivery plans include quality improvement activity
- 39 Local and national priorities from NSFs and HImPs are used to priorities service development
- 40 NSF implementation is integrated with business planning and quality improvement programmes
- 41 External guidelines are critically appraised before local adoption or development
- 43 The organisation shares a common vision for clinical governance
- 44 Leadership skills are identified and developed through leadership training programmes
- 45 Clinical teams receive performance feedback
- 47 Staff have clear and shared objectives
- 48 There are clear processes for involving service users in service development
- 49 There are clear criteria for establishing user involvement groups
- 50 Service improvement activity focuses on the patient experience of care
- 51 Local health and social care agencies work jointly on clinical governance issues
- 52 Local partnerships with health and social care agencies have clear, shared purposes
- 54 Staff evaluate the best ways of training and learning from experience

'Outcome' items

- 3 Where appropriate, staff modify their care processes to reflect the best practice of others
- 6 Research information is used consistently to inform our approach to quality improvement
- 13 Following identification of a problem from risk data, clinical quality is improved
- 18 Following identification of a problem from adverse events data, clinical quality is improved
- 21 Following identification of a problem from complaints data, clinical quality is improved
- 25 Clinical indicators are used to reflect on, review and integrate services
- 28 New skills gained through development activity are used in clinical settings
- 38 Decisions about service developments or cutbacks are made on clear criteria
- 46 Clinical teams respond to changes in their environment by reorganising their work processes
- 53 Staff question what they are doing, and are able to develop new and innovative models of service

Appendix 14: Step-down F analyses of trust type, CHI visit and board membership, using OPGC domain scores as independent variables

The table below shows results of a step-down F-test analysis of the relationships between trust type, CHI visit and board membership, using OPGC domain scores as independent variables. Order of entry was statistically derived, based on primacy of univariate F significance. It is clear that a number of variables that were non-significant in univariate F analysis became significant when the effects of covariation were controlled for. This makes analysis extremely difficult, especially where there is no compelling case for order of entry.

IV	DV	Univariate F	DF	Step-down F	DF
TYPE	IQ	6.064a	2	12.182**	2
	MR	.492	2	13.175**	2
	IP	9.148a	2	9.148**	2
	CA	7.521a	2	.195	2
	LC	2.560	2	35.000**	2
CHID	IQ	1.594	2	.823	2
	MR	.474	2	.900	2
	IP	7.374a	2	7.374**	2
	CA	3.80	2	.3089	2
	LC	.793	2	4.718**	2
BOARDMEM	IQ	6.661a	1	..398	1
	MR	8.509a	1	.095	1
	IP	.270	1	41.542**	1
	CA	16.351a	1	16.351**	1
	LC	10.122a	1	5.408	1

(a)= significance level of $p < .01$ in univariate context

** $p < .01$

Appendix 15: Publications derived from the thesis

1. Freeman T. Using performance indicators to improve health care quality in the public sector: a review of the literature. *Health Services Management Research* 2002;**15**:126-137
2. Freeman T. Measuring progress in clinical governance: assessing the reliability and validity of the Clinical Governance Climate Questionnaire (CGCQ). *Health Services Management Research* 2003;**16**:234-250
3. Freeman T. A critical introduction to clinical governance. In Sallah D and Clark M. (Eds.) *Research and development in mental health*. Oxford: Elsevier (in press).

Bibliography

1. 6, Perri. *New Labour's project for modernising public services*. Presentation to the annual congress of the Centre for Mental Health Services Development, Shrigley Hall, Cheshire, July 6th 2000
2. Adab P, Rouse AM, Mohammed MA and Marshall T. Performance league tables: the NHS deserves better. *British Medical Journal* 2002;**324**:95-98
3. Al-Assaf AF. Healthcare outcomes management and quality improvement. *Journal of the Royal Society of Health* 1996;**116**(4):245-251
4. Amabile TM. A model of creativity and innovation in organizations. *Research in Organizational Behavior* 1988;**10**:123-167
5. Anthony PD. The paradox of the management of culture or "He who leads is lost". *Personnel Review* 1990;**19**(4):3-8
6. Appleby J and Coote A. *Five-year health check: a review of government health policy 1997-2002*. London: King's Fund, 2002
7. Ashton CM and Wray NP. A conceptual framework for the study of early admission as an indicator of quality of care. *Social Science and Medicine* 1996;**43**(11):1533-1541
8. Atkinson JM and Heritage J. *Structures of social action: studies in conversation analysis*. Cambridge: Cambridge University Press, 1984
9. Audit Commission. *Aiming to improve the principles of performance measurement*. London: Audit Commission 2000

10. Audit Commission. *On target: the practice of performance indicators*. London: Audit Commission 2000
11. Ball A, Broadbent J and Moore C. Best Value and the control of local government: challenges and contradictions. *Public Money and Management* 2002;**22**:9-15
12. Barbour RS and Kitzinger J. *Developing focus group research: politics, theory and practice*. London: Sage, 1999
13. Barnetson B and Cutright M. Performance indicators as conceptual technologies. *Higher Education* 2000;**40**:277-292
14. Barnsley J, Lemieux-Charles L and Baker GR. Selecting clinical outcome indicators for monitoring quality of care. *Healthcare Management Forum* 1996;**9**(1):5-12
15. Barry J, Proops J. Seeking sustainability discourses with Q methodology. *Ecological Economics*. 1999;**28**:337-345
16. Bartlett MS. A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society* 1954;**16**(Series B):296-298
17. Bate P. *Strategies for cultural change* Oxford: Butterworth-Heinemann, 1994
18. Bernick CL. When your culture needs a makeover. *Harvard Business Review* 2001;**79**:53-61
19. Billings JR. Community development: a critical review of approaches to evaluation. *Journal of Advanced Nursing* 2002;**31**(2):472-480
20. Binleys directory of NHS Managers. Corringham: Beechwood House Publishing,

2002

21. Birch S and Maynard A. Performance indicators and performance assessment in the UK National Health Service: implications for management and planning.
International Journal of Health Planning and management 1986;**1**:143-156
22. Black N. Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal* 1996;**312**:1215-1218
23. Blalock AB. Evaluation research and the performance management movement: from estrangement to useful integration. *Evaluation* 1999;**5**(2):117-149
24. Bland JM and Altman DG. Chronbach's alpha. *British Medical Journal* 1997;**314**:572
25. Block J. *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press 1978
26. Blumberg M. Risk adjusting health care outcomes: A methodologic review. *Medical Care Review* 1986;**43**(2):351-393
27. Blumenthal D and Kilo C. A report card on continuous quality improvement. *Milbank Quarterly* 1998;**76**:625-648
28. Boland T and Fowler A. A systems perspective of performance management in public sector organisations. *The International Journal of Public Sector Management* 2000;**13**(5):417-446
29. Bourne M and Ezzamel M. Organisational culture in the National Health Service. *Financial Accountability and Management* 1986;**2**:203-223
30. Bowling A. Measuring health. *A review of quality of life measurement scales*.

Buckingham: Open University Press, 1997 (2nd Edition)

31. Boyne G. External regulation and best value in local government. *Public Money and Management* 2000;**20**:7-12
32. Broadbent J. Comprehensive performance assessment: the crock of gold at the end of the performance rainbow? *Public Money and Management* 2003;**23**:5-7
33. Brouwer M. Q is accounting for tastes. *Journal of Advertising Research*. 1999;**39**(2):35-39
34. Brown RB, McCartney S and Bell L. Why the NHS should abandon the search for the universal outcome measure. *Health Care Analysis* 1995;**3**:191-195
35. Brown S. *Political subjectivity*. New Haven: Yale University Press 1980
36. Brown SR. Q-methodology and qualitative research. *Qualitative Health Research*. 1996;**6**(4):561-567
37. Bruner GC and Hensel PJ. Multi-scale usage in marketing journals: 1980 to 1989. *Journal of the Academy of Marketing Science*. 1993;**21**:339-344
38. Bryman A and Cramer D. *Quantitative data analysis with SPSS Release 8 for Windows: a guide for social scientists*. London: Routledge, 1999
39. Buetow SA and Roland M. Clinical governance: bridging the gap between managerial and clinical approaches to quality of care. *Quality and Safety in Health Care* 1999;**8**:184-190
40. Bull AR, Hatton P, Bensley DC, Bull SJ and Fryers PT. Standardised mortality from cervical cancer: a measure of performance? *Journal of Public Health Medicine* 1994;**16**(1):16-22

41. Camp RC. *Benchmarking: the search for industry best practices that lead to superior performance*. Milwaukee, WI: Quality Press, 1989
42. Campbell S, Roland M and Leese B. Progress in clinical governance: findings from the first NPCRDC national tracker survey of Primary Care Groups / Trusts. *British Journal of Clinical Governance* 2001;6(2):90-93
43. Campbell SM, Roland MO, Quayle JA, Buetow SA and Shekele PG. Quality indicators in general practice: which ones can general practitioners and health authority managers agree are important and how useful are they? *Journal of Public Health Medicine* 1998;20:414-421
44. Campbell S, Roland M and Wilkin D. Improving the quality of care through clinical governance. *British Medical Journal* 2001;322:1580-1582 (a)
45. Campbell SM, Sheaff R, Sibbald B, Marshall MN, Pickard S, Gask L, Halliwell S, Rogers A and Roland MO. Implementing clinical governance in English Primary Care Groups / Trusts: reconciling quality improvement and quality assurance. *Quality and Safety in Health Care* 2002;11:9-14
46. Carey MA. Comment: concerns in the analysis of focus group data. *Qualitative Health Research* 1995;5(4):487-495
47. Carter N. Performance indicators: backseat driving or hands off control? *Policy and Politics* 1989;17(2):131-138
48. Carter N, Klein R and Day P. *How organizations measure success: the use of performance indicators in government*. London: Routledge 1992
49. Cattell RB. *Handbook of multivariate experimental psychology*. Chicago: Rand

McNally, 1966

50. Cattell RB. *The scientific uses of factor analysis in behavioural and life sciences*. New York: Plenum, 1978
51. Cave M, Kogan M and Smith R. *Output and performance measurement in government: the state of the art*. London: Jessica Kingsley Publishers 1990
52. Challiner Y, Carpenter GI, Potter J and Maxwell C. Performance indicators for hospital services for older people. *Age and Ageing* 2003;**32**(3):343-346
53. Chen TS. *Theory driven evaluation*. London: Sage, 1990
54. Child D. *The essentials of factor analysis*. London: Holt Reinhart and Winston, 1976
55. Chinnis AS, Paulson DJ and Davies SM. Using Q-methodology to assess the needs of emergency medicine support staff employees. *Journal of Emergency Medicine*. 2001;**20**(2):197-203
56. Clarke A and Dawson R. *Evaluation research*. London: Sage, 1999
57. Clarke J, Gewirtz S and McLaughlin E (eds.) *New managerialism, new welfare?* London: Sage, 2000
58. Clarke J and Newman J. *The managerial state* London: Sage, 1997
59. Clarkson P. and Challis D. Developing performance indicators for mental health care. *Journal of Mental Health* 2002;**11**(3):281-293
60. Clinical Staff Questionnaire: draft paper. Commission for Health Improvement (unpublished communication), 2000
61. Collopy BT. Health care performance measurement systems and the ACHS care

- evaluation program. *Journal of Quality in Clinical Practice*. 1998;**18**:171-176
62. Comrey AL and Lee HB. *A first course in factor analysis*. (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates, 1992
63. Conduit D, Morgan A and Willetts J. Clinical Governance in NHS trusts: results of an early baseline assessment in the Trent region. *The Journal of Clinical Governance* 1999;**7**:121-123
64. Cook TD and Reichardt CS (eds.) *Qualitative and quantitative methods in evaluation research* London: Sage, 1979
65. Cowpe J and Scrivens E. *The impact of clinical governance on medical and clinical directors*. Keele: University of Keele, 2001
66. Coyle IR, Sleeman SD, Adams N. Safety climate. *Journal of Safety Research* 1995;**26**(4):247-254
67. Coyle YM and Battles JB. Using antecedents of medical care to develop valid quality measures. *International Journal for Quality in Health Care* 1999;**11**(1):5-12
68. Cramer D. *Fundamental statistics for social research. Step by step calculations and computer techniques using SPSS for Windows*. London: Routledge, 1998
69. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297-334
70. Cullen R, Nicholls S and Halligan A. Reviewing a service – discovering the unwritten rules. *British Journal of Clinical Governance* 2000;**5**(4):233-239
71. Davenport RJ, Dennis MS and Warlow CP. Effect of correcting outcome data for

- case mix: an example from stroke medicine. *British Medical Journal* 1996;**312**:1503-1505
72. Davies HTO. Performance management using health outcomes: in search of instrumentality. *Journal of Evaluation in Clinical Practice* 1998;**4**(4):359-362
73. Davies HTO. Health service culture and health service performance. *British Journal of Health Care Management* 2000;**6**(12):558-560
74. Davies HTO and Crombie IK. Assessing the quality of care. *British Medical Journal* 1995;**311**:766
75. Davies HTO and Crombie IK. Interpreting health outcomes. *Journal of Evaluation in Clinical Practice* 1997;**3**(3):187-199
76. Davies HTO and Lampel J. Trust in performance indicators? *Quality in Health Care* 1998;**7**:159-162
77. Davies HTO and Mannion R. *Clinical governance: striking a balance between checking and trusting*. York: Centre for Health Economics, University of York , 1999a
78. Davies HTO and Mannion R. The rise of oversight and the decline of mutuality? *Public Money and Management* 1999b;**19**:55-59
79. Davies HTO and Mannion R. *Clinical governance: striking a balance between checking and trusting*, in Smith PC (ed.) *Reforming markets in health care: an economic perspective*. Buckingham: Open University Press, 2000
80. Davies HTO, Mannion R and Marshall MN. Treading a third way for quality in health care. *Public Money and Management* 2001;**21**:6-7

81. Davies HTO and Nutley SM. Developing learning organisations in the NHS.
British Medical Journal 2000;**320**:998-1001
82. Davies HTO Nutley S and Mannion R. Organisational culture and quality of health care. *Quality and Safety in Health Care* 2000;**9**:111-119
83. Davies JS. The governance of urban regeneration: A critique of the 'governing without government' thesis. *Public Administration* 2002;**80**(2):301-322
84. Dawson S. *Managing, organising and performing in health care: what do we know and how can we learn?* In Mark A and Dopson S (eds.) *Organisational behaviour in health care*. London: Macmillan, 1999
85. Day JC, Bentall RP and Warner S. Schizophrenic patients' experiences of Neuroleptic medication: A Q-methodological investigation. *Acta Psychiatrica Scandinavica*. 1996;**93**(5):397-402
86. Day P and Klein R. Commission for Health Improvement invents itself. *British Medical Journal* 2001;**322**:1502-1503
87. Day P and Klein R. Who nose best? *Health Services Journal* 2002;**112**(5799):26-29
88. Deeming C. Performance indicators: making information out of data. *Health Care UK* 2001;(spring):33-41
89. Degeling P, Hunter DJ and Dowswell B. Changing health care systems. *Journal of Integrated Care Pathways* 2001;**5**:64-69
90. Degeling P, Kennedy J, Hill M, Carnegie M and Holt J. *Professional sub-cultures and hospital reform*. New South Wales: University of New South Wales, 1999

91. Degeling P, Kennedy J and Hill M. Mediating the cultural boundaries between medicine, nursing and management- the central challenge in hospital reform. *Health Services Management Research* 2001;**14**:36-48
92. Degeling P, Maxwell S, Kennedy J and Coyle B. Medicine, management, and modernisation: a “danse macabre”? *British Medical Journal* 2003;**326**:649-652
93. Dence R. Best practices in benchmarking. In Holloway J, Lewis J and Mallory G. (eds.) *Performance measurement and evaluation*. London: Sage, 1995
94. Denzin NK. The logic of naturalistic inquiry. In NK Denzin (ed.) *Sociological methods: a sourcebook*. New York: McGraw-Hill, 1978
95. Department of Health. *The new NHS: modern, dependable*. London: HMSO, 1997
96. Department of Health. *A First Class Service: Quality in the new NHS*. London: HMSO, 1998
97. Department of Health. *Clinical governance: quality in the new NHS*. London: HMSO, 1999a
98. Department of Health. *Health Act 1999*. London: HMSO, 1999b
99. Department of Health *National service framework for mental health services: modern standards & service models* London: HMSO, 1999c
100. Department of Health *An organisation with a Memory* London: HMSO, 2000a
101. Department of Health. *Clinical governance in London trusts: Taking stock: the 1999 London Region stocktake of clinical governance: An overview of activities across all London trusts*. London: NHS London Regional Office, 2000b
102. Department of Health *National service framework for coronary heart disease:*

- modern standards & service models* London: HMSO, 2000c
103. Department of Health *The NHS Plan* London: HMSO, 2000d
 104. Department of Health *Quality and Performance in the NHS Performance Indicators* London: HMSO, 2000e
 105. Department of Health *Assuring the quality of medical practice* London: HMSO, 2001a
 106. Department of Health. *Draft research governance framework for health and social care.* London: The Stationery Office, 2001b
 107. Department of Health *National service framework for older people: modern standards & service models* London: HMSO, 2001c
 108. Department of Health *NHS Performance Ratings: Acute Trusts 200/01* London: HMSO, 2001d
 109. Department of Health. *Shifting the balance of power.* London: HMSO, 2001e
 110. Department of Health. *Delivering the NHS Plan (Cm5503).* London: HMSO, 2002a
 111. Department of Health *National service framework for diabetes: delivery strategy* London: HMSO, 2002b
 112. Department of Health. *NHS Foundation trusts.* London: HMSO, 2002c
 113. Department of Health *NHS Reform and Healthcare Professions Act* London: HMSO, 2002d
 114. Department of Health *National service framework for diabetes: modern standards & service models* London: HMSO, 2003

115. DeVellis RF. *Scale development: theory and applications*. Newbury Park CA: Sage, 1992
116. Dewar S. *Clinical governance under construction: Problems of design and difficulties in practice*. London: King's Fund, 1999
117. Diamond CC, Rask KJ and Kohler SA. Use of paper medical records versus administrative data for measuring and improving health care quality: are we still searching for a gold standard? *Disease Management* 2001;**4**(3):121-130
118. Donabedian A. The role of outcomes in quality assessment and assurance. *Quality Review Bulletin* 1992;**18**(11):356-360
119. Drummond MF, O'Brien B, Stoddart GL and Torrance W. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, 2nd edition 1997
120. Edwards N, Marshall M, McLellan A and Abbasi K. Doctors and managers: a problem without a solution? *British Medical Journal* 2003;**326**:609-610
121. EFQM *The EFQM Excellence Model: public and voluntary sector* Brussels: The European Foundation for Quality Management, 1999
122. Ekvall G, Ryhammar L. The creative climate: its determinants and effects at a Swedish university. *Creativity Research Journal* 1999;**12**(4):303-310
123. Ekvall G. *Climate, structure and innovativeness of organisations*. Working paper of the Swedish Council for Management and Organisational Behaviour, 1983
124. Emsley S. Governance and internal control. *Health Care Risk Report* 2002;**6**:20-21

125. Ferguson B and Lim JNW. Incentives and clinical governance: money following quality? *Journal of Management in Medicine* 2001;**15**(6):463-487
126. Ferlie EB and Shortell SM. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *The Milbank Quarterly* 2001;**79**(2):281-315
127. Firth-Cozens J. Celebrating teamwork. *Quality in Health Care* 1998;**7**(Suppl):S3-S7
128. Flynn N. Performance measurement in public sector services. *Policy and Politics* 1986;**14**(3):389-404
129. Forster DP, Frost CEB and Morris D. Modifying practice: Two initiatives in the English National Health Service. *Journal of Public Health Policy* 1990;**11**(1):81-105
130. Franks A. How goes the night, watchman? An overview of the first annual clinical governance reports (1999/2000) from acute trusts in an English NHS region. *Journal of Management in Medicine* 2001;**15**(3):220-226
131. Freeman T. Learning from the West Midlands Regional Office clinical governance visit programme. *The Journal of Clinical Governance*. 2000;**8**(4):107-110
132. Freeman T. Using performance indicators to improve health care quality in the public sector: a review of the literature. *Health Services Management Research* 2002;**15**:126-137
133. Freeman T, Latham L, Walshe K, Spurgeon P and Wallace L. *The early*

development of clinical governance: a survey of NHS trusts in the Southwest region.

Project Report 12. Birmingham: HSMC, University of Birmingham, 1999

134. Freeman T, Latham L, Walshe K, Wallace L and Spurgeon P. Evaluating mandated quality improvement: results of a survey in the NHS in England. *Journal of Clinical Excellence* 2001;**3**(1):15-22 (a)
135. Freeman T, Latham L, Walshe K, Wallace L and Spurgeon P. How do trusts intend to measure progress in clinical governance? *The Journal of Clinical Governance* 2001;**9**(1):37-43(b)
136. Garside P. Organizational context for quality: lessons from the fields of organizational development and change management. *Quality in Health Care* 1998;**7**(Suppl):S8-S15
137. Geddes M and Martin S. The policy and politics of Best Value: currents, crosscurrents and undercurrents in the new regime. *Policy and Politics* 2000;**28**(3):379-395
138. Gerowitz MB. Do TQM interventions change management culture? Findings and implications. *Quality Management in Health Care* 1998;**6**(3):1-11
139. Giuffrida A and Gravelle H. Measuring performance in primary care: econometric analysis and DEA. *Applied Economics* 2001;**33**(2):163-175
140. Giuffrida A, Gravelle H and Roland M. Measuring quality of care with routine data: avoiding confusion between performance indicators and health outcomes. *British Medical Journal* 1999;**319**:94-98
141. Giuffrida A, Gravelle H and Roland M. *Performance indicators for managing*

- primary care: The confounding problem.* In: Smith P. (ed.) *Reforming markets in health care: an economic perspective.* Buckingham: Open University Press 2000:162-186
142. Glick WH. Conceptualising and measuring organisational and psychological climate: pitfalls in multilevel research. *Academy of Management Review* 1985;**10**(3):601-616
143. Goddard M, Mannion R and Smith PC. Assessing the performance of NHS hospital trusts: the role of 'hard' and 'soft' data. *Health Policy* 1999;**48**(2):119-134
144. Goddard M, Mannion R and Smith P. Enhancing performance in health care: a theoretical perspective on agency and the role of information. *Health Economics.* 2000;**9**(2):95-107
145. Goddard M, Mannion R and Smith P. *The performance framework: taking account of economic behaviour.* In: Smith P. (ed.) *Reforming markets in health care: an economic perspective.* Buckingham: Open University Press, 2000
146. Goldstein H and Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society A* 1996;**159**:385-443
147. Goodman N. Clinical governance. *British Medical Journal* 1998;**317**:1725-1727
148. Gravelle H, Smith P and Xavier A. Performance signals in the public sector: the case of health care. *Oxford Economic Papers* 2003;**55**(1):81-103
149. Greene JC. The inequality of performance measurements. *Evaluation*

1999;5(2):160-172

150. Greener I. Performance in the National Health Service: the insistence of measurement and the confusion of content. *Public Performance and Management Review* 2003;26(3):237-250
151. Gross PA, Braun B, Kritchevsky SB and Simmons BP. Comparison of clinical indicators for performance measurement of health care quality: a cautionary note. *British Journal of Clinical Governance* 2000;5(4):202-211
152. Guadagnoli E and Velicer WF. Relation of sample size and the stability of component patterns. *Psychological Bulletin*. 1988;103:265-275
153. Guba E and Lincoln E. *Fourth Generation Evaluation* New York: Sage, 1989
154. Guldenmund FW. The nature of safety culture: a review of theory and research. *Safety Science* 2000;34:215-257
155. Guzzo RA, Dickson MW. Teams in organizations: recent research on performance and effectiveness. *Annual Review of Psychology* 1996;47:307-338
156. Hackett M, Lilford R, Jordan J. Clinical governance: culture, leadership and power- the key to changing attitudes and behaviours in trusts. *International Journal of Health Care Quality* 1999;12:98-104
157. Hager P and Gonczi A. What is competence? *Medical Teacher* 1996;18(1);15-18
158. Halligan A and Donaldson L. Implementing clinical governance: turning vision into reality. *British Medical Journal* 2001;322:1413-1417
159. Ham C. Improving NHS performance: human behaviour and health policy.

- British Medical Journal* 1999a;**319**:1490-1492
160. Ham C. The third way in health reform: does the emperor have any clothes?
Journal of Health Services Research and Policy 1999b;**4**(3):168-173
161. Hannan EH. Measuring hospital outcomes: don't make perfect the enemy of the good! *Journal of Health Service Research and Policy* 1998;**3**:67-69
162. Hannan EL, Kilburn H, Racz M, Shields E and Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. *Journal of the American Medical Association* 1994;**271**(10):761-766
163. Hargreaves S. Three trusts deliberately misreported data, says Audit Commission. *British Medical Journal* 2003;**326**:517
164. Harley MJ. Performance indicators in the management process. *Health Services Management Research* 1988;**1**(1):29-42
165. Harley M and Yates J. Performance indicators in the health service. *Public Finance and Accountancy* 1987; Nov 6th :9-12
166. Harrison S and Dowswell G. *The selective use by NHS management of NICE-promulgated guidelines: a new and effective tool for systematic rationing of new therapies?* In Miles A, Hampton JR and Hurwitz B. (eds.) *NICE, CHI and the NHS reforms: enabling excellence or imposing control?* London: Aesculapius Press, 2000
167. Harrison S and Lim JNW. Clinical governance and primary care in the English National Health Service: some issues of organisation and rules. *Critical Public Health* 2000;**10**(3):321-32

168. Harrison S and Lim NW. The frontier of control: doctors and managers in the NHS 1966 to 1997. *Clinical Governance: an International Journal*. 2003;**8**(1):13-18
169. Harrison S, Hunter D, Marnoch G, Pollitt C. *Just managing: power and culture in the National Health Service*. London: Macmillan, 1992
170. Hauck K, Rice N and Smith P. The influence of health care organisations on health system performance. *Journal of Health Services Research and Policy* 2003;**8**(2):68-74
171. Hawkins P. Organizational culture: sailing between evangelism and complexity. *Human Relations* 1997;**50**(4):417-441
172. Heasel S. The cost of performance indicators. *Health Care Risk Report* 2000;**6**(4):21-22
173. Heaton C. External Peer review in Europe: an overview from the EXPeRT project. *International Journal for Quality in Health Care*. 2000;**12**(3):177-182
174. Helms M and Stern R. Exploring the factors that influence employees' perceptions of their organisation's culture. *Journal of Management in Medicine* 2001;**15**(6):415-429
175. Hendrix MS, Wakefield DS, Murray JF, Uden-Holman T, Helms CM and Ludke RL. Using comparative clinical and economic outcome information to profile physician performance. *Health Services Management Research* 1995;**8**(4):213-220
176. Hepworth NP. Measuring performance in non-market organizations. *International Journal of Public Sector Management* 1988;**1**(1):16-26

177. Hertz H, Reinman C, Bostwick M. The Malcolm Baldrige national quality award concept: could it help stimulate or accelerate healthcare quality improvement? *Quality Management in Health Care* 1994;**2**:63-72
178. Hofer TP and Hayward RA. Identifying poor quality hospitals: can hospital mortality rates detect quality problems for medical diagnoses? *Medical Care* 1996;**34**(8):737-753
179. Hofstede GR. *Cultures and organisations: software of the mind*. London: McGraw-Hill, 1991
180. Hofstede GR, Neuijen B, Ohayv DD and Sanders G. Measuring organisational cultures: a qualitative and quantitative study across twenty sites. *Administrative Science Quarterly* 1990;**35**:286-316
181. Hoggett P. New modes of control in the public service. *Public Administration* 1996;**74**:9-32
182. Holland K and Fennell S. Clinical governance is 'ACE' - using the EFQM excellence model to support baseline assessment. *International Journal of Health Care Quality Assurance* 2000;**13**(4):170-177
183. Holland WW, Fitzgerald A.P, Hildrey SJ and Phillips SJ. Heaven can wait. *Journal of Public Health Medicine* 1994;**16**(3):321-330
184. Holloway J, Francis G and Hinton M. A vehicle for change? A case study of performance improvement in the 'new' public sector. *The International Journal of Public Sector Management* 1999;**12**(4):351-365
185. Hood C, James O, Jones G, Scott C and Travers T. Regulation inside

- government: where the new public management meets the audit explosion. *Public Money and Management* 1998;**18**(2):61-68
186. Hood C, Scott C, James O, Jones G and Travers T (eds.) *Regulation inside government: waste watchers, quality police and sleaze busters*. Oxford: Oxford University Press, 1999
187. Horwood J. The search for diversity in male identity using Q-methodology. *Feminism and Psychology*. 2000;**10**(4):492-497
188. Houghton G, Wall D and Field S. Primary care group clinical governance leads: their educational needs and experiences. *Journal of Clinical Governance* 1999;**7**:190-194
189. Ibrahim JE. Performance indicators from all perspectives. *International Journal for Quality in Health Care* 2001;**13**(6):431-432
190. Iezzoni LI, Swartz M, Ash AS, Hughes JS, Daley J and Mackiernan YD. Using severity-adjusted stroke mortality rates to judge hospitals. *International Journal for Quality in Health Care* 1995;**7**(2):81-94
191. Iles V and Sutherland K. *Organisational change: a review for health care managers, professionals and researchers*. London: National Co-ordinating Centre for NHS Service Delivery and Organisation R&D, 2000
192. Isaksen SG, Lauer KJ, Ekvall G. Situational outlook questionnaire: a measure of the climate for creativity and change. *Psychological Reports* 1999;**85**:665-674
193. Jackson A. An evaluation of evaluation: problems with performance measurement in small business loan and grant schemes. *Progress in Planning*

2001;**55**:1-64

194. Jackson S. Exploring the possible reasons why the UK government commended the EFQM excellence model as the framework for delivering governance in the new NHS. *International Journal of Health Care Quality Assurance* 1999;**12**(6):244-253
195. Jackson S. Achieving clinical governance in women's services through the use of the EFQM excellence model. *International Journal of Health Care Quality Assurance* 2000;**13**(4):182-190
196. Jacobs K and Manzi T. Performance indicators and social constructivism: conflict and control in housing management. *Critical Social Policy* 2000;**20**(1):85-103
197. James LR and Jones AP. Organisational climate: a review of theory and research. *Psychological Bulletin* 1974;**81**(12):1096-1112
198. Jenkins R. Towards a system of outcome indicators for mental health care. *British Journal of Psychiatry*. 1990;**157**:500-514
199. Jobanputra R and Buchan J. Power sharing. *Heath Service Journal* 2003;**13**:(5853):26-27
200. Jones CS. Towards benchmarking in British acute hospitals. *Health Services Management Research*. 2001;**14**:125-138
201. Jones GT. Indicator based systems of performance management in the National Health Service: a comparison of the perceptions of local and national level managers. *Health Services Management Research* 2000;**13**:16-26
202. Kaiser HF. An index of factorial simplicity. *Psychometrika* 1974;**39**:31-36

203. Karim K. Q-methodology- advantages and the disadvantages of this research method. *Journal of Community Nursing* 2001;15(4):8-10
204. Katzenbach JR, Smith DK. The Discipline of Teams. *Harvard Business Review* 1993;71:111-120
205. Kazandjian V, Woop P and Lawther SJ. Balancing science and practice in indicator development. *International Journal for Quality in Healthcare* 1998;7:39-46
206. Kazandjian VA, Thomson RG, Law WR and Waldron K. Do performance indicators make a difference? *The Joint Commission Journal of Quality Improvement* 1996;22(7):482-491
207. Keeton K and Mengistu B. The perception of organisational culture by management level: implications for training and development. *Public Productivity and Management Review* 1992;16(2):205-213
208. Kendrick S, Cline D and Finlayson A. *Clinical outcomes indicators in Scotland: lessons and prospect*. In Davies HTO, Tavakoli M, Malek M and Neilson AR (eds.) *Managing quality: strategic issues in health care management*. Aldershot: Ashgate Publishing Ltd, 1999
209. Kim JO and Mueller CW. *Introduction to factor analysis: what it is and how to do it*. London: Sage, 1978
210. King L and Appleton JV. Fourth generation evaluation of health services: exploring a methodology that offers equal voice to consumer and professional stakeholders. *Qualitative Health Research* 1999;9(5):698-710

211. Kitzinger C. *Introducing and developing Q as a feminist methodology: a study of accounts of lesbianism*. In Wilkinson S. (ed.) *Feminist Social Psychology: developing theory and practice*. Milton Keynes: Open University Press, 1986
212. Kitzinger J. The methodology of focus group interviews: the importance of interaction between research participants. *Sociology of Health and Illness* 1994;**16**:103-121
213. Kitzinger J. Qualitative research: introducing focus groups. *British Medical Journal* 1995;**311**:299-302
214. Kitzinger J. *Focus groups with users and providers of health care*. In Pope C and Mays N (eds.) *Qualitative research in health care*. London: BMJ Publishers (2nd edition), 2000
215. Kivimaki M, Elovainio M. A short version of the Team Climate Inventory: Development and psychometric properties. *Journal of Occupational and Organizational Psychology* 1999;**72**:241-246
216. Kivimaki M, Kuk G, Elovainio M, Thomson L, Kalliomaki-Levanto T, Heikkila A. The Team Climate Inventory (TCI)- four of five factors? Testing the structure of TCI in samples of low and high complexity jobs. *Journal of Occupational and Organizational Psychology* 1997;**70**:375-389
217. Klazinga N. Re-engineering trust: the adoption and adaption of four models for external quality assurance of health care services in western European health care systems. *International Journal for Quality in Health Care*. 2000;**12**(3):183-189
218. Klein P. *An easy guide to factor analysis*. London: Routledge, 1994

219. Klijn EH and Koppenjan JFM. Public management and policy networks: Foundations of a network mode of governance. *Public Management* 2000;2(2):135-158
220. Knodel J. *The design and analysis of focus group studies: a practical approach*. In Morgan D, ed. *Successful focus groups: advancing the state of the art*. London: Sage, 1993:35-50
221. Koivula M, Paunonen M and Laippala P. Burnout among nursing staff in two Finnish hospitals. *Journal of Nursing Management* 2000;8:149-158
222. Kooiman J. *Findings, speculations and recommendations*, in Kooiman J. (ed.) *Modern governance*. London: Sage, 1993
223. Kouzmin A, Loffler E, Klages H and Korac-Kakabadse N. Benchmarking and performance measurement in public sectors: Towards learning for agency effectiveness. *The International Journal of Public Sector Management* 1999;12(2):121-144
224. Krivenko C and Chodroff C. The analysis of clinical outcomes: getting started in benchmarking. *The Joint Commission Journal on Quality Improvement* 1994;20(5):260-266
225. Krueger RA. *Focus groups: a practical guide for applied research*. Thousand Oaks, CA: Sage, 1988
226. Krueger RA. *Quality control in focus groups*. In Morgan D (ed.) *Successful focus groups: advancing the state of the art*. London: Sage, 1993:65-85
227. Latham L, Freeman T, Walshe K, Spurgeon P and Wallace L. *The early*

- development of clinical governance: a survey of NHS trusts in the West Midlands.*
- Project Report 11. Birmingham: HSMC, University of Birmingham, 1999
228. Latham L, Freeman T, Walshe K, Spurgeon P and Wallace L. Clinical governance in the West Midlands and South West regions: early progress in NHS trusts. *Clinician in Management* 2000;9:83-91
229. Lawton A. McKevitt D and Millar M. Coping with ambiguity: reconciling external legitimacy and organisational implementation in performance measurement. *Public Money and Management* 2000;20:13-19
230. Lee R. Clinical governance and risk management. *Journal of the Medical Defence Union*. 1999;15(2):9-12
231. Leggat SG, Narine L, Lemieux-Charles L, Barnsley J, Baker GR, Sicotte C, Champagne F and Bilodeau H. A review of organizational performance assessment in health care. *Health Services Management Research* 1998;11:3-23
232. Legge K. *Managing culture: fact or fiction?* In Sisson K. (ed.) *Personnel Management*. Oxford: Blackwell, 1994
233. Leng GC, Walsh D, Fowkes FGR and Swainson CP. Is the emergency readmission rate a valid outcome indicator? *Quality in Health Care* 1999;8:234-238
234. Leyland AH and Boddy FA. Measuring performance in hospital care: Length of stay in Gynaecology. *European Journal of Public Health* 1997;7:136-143
235. Leyland AH, Pritchard CW, McLoone P and Boddy FA. Measures of performance in Scottish maternity hospitals. *British Medical Journal* 1991;303:389-

236. Linder JC. Outcomes measurement: compliance tool or strategic initiative?
Health Care Management Review 1991;**16**(4):21-33
237. Logan S. Outcome measures in child health. *Archives of Disease in Childhood*
1991;**66**:745-748
238. Lorence DP, Spink A and Jameson R. Information in medical decision-making:
how consistent is our management? *Medical Decision-Making* 2002;**22**(6):514-521
239. Lowndes V and Skelcher C. The dynamics of multi-organizational partnerships:
an analysis of changing modes of governance. *Public Administration* 1998;**76**:313-
333
240. Lowry S. Focus on performance indicators. *British Medical Journal*
1988;**296**:992-994
241. Mannion R, Davies HTO, Marshall MN. *Cultures for performance in health
care: evidence on the relationship between organisational culture and
organisational performance in the NHS*. York: University of York, 2003a
242. Mannion R, Davies HTO, Marshall MN. *Impact of 'star' performance ratings
on English acute trusts: secondary analysis of case-study evidence*. London:
Commission for Health Improvement, 2003
243. Mannion R and Goddard M. Impact of published clinical outcomes data: case
study in NHS hospital trusts. *British Medical Journal* 2001;**323**:260-263
244. Mant J. Detecting differences in quality of care: the sensitivity of measures of
process and outcome in treating acute myocardial infarction. *British Medical*

- Journal* 1995;**311**:793-796
245. Mant J and Hicks N. Health status measurement and the assessment of medical care. *International Journal for Quality in Health Care* 1996;**8**(2):107-109
 246. Marshall M. Accountability and quality improvement: the role of report cards. *Quality in Health Care* 2001;**10**:67-68
 247. Marshall M, Sheaff R, Rogers A, Campbell S, Halliwell S, Pickard S, Sibbald B, Roland M. A qualitative study of the cultural changes in primary care organisations needed to implement clinical governance. *British Journal of General Practice* 2002;**52**: 641-645
 248. Martin J. *Organizational culture: mapping the terrain*. Thousand Oaks, CA: Sage, 2002
 249. Martin LL and Kettner PM. Performance measurement: The new accountability. *Administration in Social Work* 1997;**21**(1):17-29
 250. Martin V. Service planning and governance. *Nursing Management* 2001;**8**(2):32-36
 251. Mayntz R. *Governing failures and the problems of governability: some comments on a theoretical paradigm*. In Kooiman J. (ed.) *Modern governance*. London: Sage, 1993
 252. McCartney S and Berman-Brown R. Managing by numbers: using outcome measures in the NHS. *International Journal of Health Care Quality Assurance* 1998;**12**(1):6-12
 253. McColl A, Gabbay J and Roderick P. Improving health outcomes- a review of

- case studies from English health authorities. *Journal of Public Health Medicine* 1998;**20**(3):302-311
254. McColl A, Roderick P, Gabbay J, Smith H and Moore M. Performance indicators for primary care groups: an evidence based approach. *British Medical Journal* 1998;**317**:1354-1360
 255. McColl A, Roderick P, Smith H, Wilkinson E, Moore M, Exworthy M and Gabbay J. Clinical governance in primary care groups: the feasibility of deriving evidence-based performance indicators. *Quality in Health Care* 2000;**9**:90-97
 256. McGlynn EA. Choosing and evaluating clinical performance measures. The *Joint Commission Journal on Quality Improvement* 1998;**24**(9):470-479
 257. McKee M. Measuring hospital performance: are we asking the right questions? *Journal of the Royal Society of Medicine* 1997;**90**:187-191
 258. McKee M. Indicators of clinical performance: problematic, but poor standards of care must be tackled. *British Medical Journal* 1997;**315**:142
 259. McKee M. Measuring performance in the NHS. *British Medical Journal* 1998;**316**:322
 260. McKee M and Hunter D. Mortality league tables: do they inform or mislead? *Quality in Health Care* 1995;**4**:5-12
 261. McKee M and James P. Using routine data to evaluate quality of care in British hospitals. *Medical Care* 1997;**35**(10) Suppl.:OS102-OS111
 262. McKenzie FC and Shilling MD. Avoiding performance measurement traps: Ensuring effective incentive design and implementation. *Compensation and*

- Benefits Review* 1998;**30**(4):57-65
263. McKeown M, Hincks M, Stowell-Smith M, Mercer D and Forster J. Q-methodology, risk training and quality management. *International Journal of Health Care Quality Assurance*. 1999;**12**(6):254-266
264. McKeown M, Stowell-Smith M and Foley B. Passivity vs. militancy: a Q-methodological study of nurses industrial relations on Merseyside (England). *Journal of Advanced Nursing*. 1999;**30**(1):140-149
265. McKeown B and Thomas D. *Q -methodology*. Newbury Park: Sage 1988
266. McKevitt D. Managing for results: the role of performance measurement in health care delivery. *Administration* 1993;**41**(3):307-322
267. McLaughlin V, Leatherman S, Fletcher M and Wyn-Owen J. Improving performance using indicators. Recent experiences in the Unites States, the United Kingdom, and Australia. *International Journal for Quality in Health Care* 2001;**13**(6):455-462
268. Meekings A. Unlocking the potential of performance measurement: A practical implementation guide. *Public Money and Management* 1995;**15**(4):5-12
269. Merton RK. *The focused interview*. Glencoe, IL: Free Press, 1956
270. Miles A, Hampton JR and Hurwitz B. *NICE, CHI and the NHS reforms: enabling excellence or imposing control?* London: Aesculapius Press, 2000
271. Miles B and Huberman AM. *Qualitative data analysis: An expanded sourcebook*. Newbury Park, CA: Sage, 1994 (2nd edition)
272. Milne R and Clarke A. Can readmission rates be used as an outcome indicator?

- British Medical Journal* 1990;**301**:1139-1140
273. Modernisation Agency. *Essence of care: patient focused benchmarks for clinical governance*. London: Department of Health, 2003
 274. Moeller J. The EFQM Excellence Model. German experiences with the EFQM approach in healthcare. *International Journal for Quality in Health Care* 2001;**13**(1):45-49
 275. Moeller J, Breinlinger-O'Reilly J and Elser J. Quality measurement in German health care- the EFQM Excellence Model. *International Journal of Health Care Quality Assurance* 2000;**13**(6):254-258
 276. Mohammed MA, Cheng KK, Rouse A and Marshall T. Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons. *The Lancet* 2001;**357**:463-467
 277. Morgan D. *Focus interviews as qualitative research*. Newbury Park, CA: Sage, 1988
 278. Morgan DL and Krueger RA. *When to use focus groups and why*. In Morgan D (ed.) *Successful focus groups: advancing the state of the art*. London: Sage, 1993: 3-19
 279. Moullin M. *Delivering excellence in health and social care*. Buckingham: Open University Press, 2002
 280. Mullen PM. Performance indicators- is anything new? *Hospital and Health Services Review* 1985;165-167
 281. Mulligan J, Appleby J and Harrison A. Measuring the performance of health systems: indicators still fail to take socio-economic factors into account. *British*

- Medical Journal* 2000;**321**:191-2
282. Murphy E, Dingwall R, Greatbach D, Parker S and Watson P. *Qualitative research methods in health technology assessment: a review of the literature*. Health Technology Assessment 1998;**2**(16)
 283. Nerenz D. Accountability for health outcomes and the proper unit of analysis: what do the experts think? *International Journal for Quality in Health Care* 1998;**10**(6):539-546
 284. Newman J. *Beyond the New Public Management? Modernising public services*. In Clarke J, Gerwitz S and McLaughlin E (eds.) *New managerialism, new welfare?* London: Sage, 2000
 285. Newman K, Pyne T, Leigh S, Rounce K, Cowling A. Personal and organizational competencies requisite for the adoption and implementation of evidence-based healthcare. *Health Services Management Research*. 2000;**13**:97-110
 286. NHS Executive West Midlands Region. *West Midlands clinical governance self-assessment & development guide*. Birmingham: NHSE, 1999
 287. NHS Executive *Clinical governance: Quality in the new NHS* London: HMSO 1999
 288. NHS Executive. *The NHS performance assessment framework*. Wetherby: Department of Health, 1999
 289. Nicholls S, Cullen R, O'Neil S and Halligan A. Clinical governance: its origins and foundations. *British Journal of Clinical Governance* 2000;**5**(3):172-178

290. Norusis MJ. *SPSS advanced statistics guide*. Chicago, Illinois: SPSS Inc, 1990
291. Nunnally J. *Psychometric theory*. New York: McGraw Hill, 1978
292. Nutley S and Smith PC. League tables for performance improvement in health care. *Journal of Health Services Research and Policy* 1998;3(1):50-57
293. Nyhan RC and Marlowe HA. Performance measurement in the public sector: Challenges and opportunities. *Public Productivity and Management Review* 1995;18(4):333-348
294. Oakland JS. *Statistical process control* Oxford: Butterworth Heinemann, 5th edn. 2003
295. O'Brien K. *Improving survey questionnaires through focus groups*. In Morgan D (ed.) *Successful focus groups: advancing the state of the art*. London: Sage, 1993:105-118
296. Ogbonna E and Harris LC. Managing organizational culture: compliance or genuine change? *British Journal of Management* 1998;9:273-288
297. Oppenheim AN. *Questionnaire design, interviewing and attitude measurement*. London: Pinter, 1992
298. Orchard C. Comparing health outcomes. *British Medical Journal* 1994;308:1493-1496
299. Osborne D and Gaebler T. *Reinventing government: how the entrepreneurial spirit is transforming the public sector*. Reading, MA: Addison-Wesley, 1992
300. Ouchi W. *Markets, bureaucracies and clans*. In Thompson G, Frances J, Levacic R and Mitchell J (eds.). *Markets, hierarchies and networks: the co-*

- ordination of social life*. London: Sage, 1991
301. Ouchi WG. A conceptual framework for the design of organizational control mechanisms. *Management Science*. 1979;**25**(9):833-848
 302. Ovreteit J. *Evaluating health interventions*. Buckingham: Open University Press, 1998
 303. Page CK. Performance improvement integration: a whole systems approach. *Journal of Nursing Care Quality* 1999;**13**(3):59-70
 304. Parry GJ, Gould CR, McCabe CJ and Tarnow-Mordi WO. Annual league tables of mortality in neonatal intensive care units: longitudinal study. *British Medical Journal* 1998;**316**:1932-1935
 305. Patton MQ. *Utilization-focused evaluation*. London: Sage, 1986
 306. Patton MQ. *How to use qualitative methods in evaluation*. London: Sage, 1987
 307. Pawson R and Tilley N. *Realistic evaluation*. London: sage, 1997
 308. Peck E. Modernising the NHS: a game of two halves. *British Journal of Health Care Management* 2001;**7**(5):198-201
 309. Pedhazur EJ and Schmelkin LP. *Measurement, design and analysis: An integrated approach*. Hillside, NJ: Erlbaum, 1991
 310. Peters T and Waterman R. *In search of excellence: lessons from America's best-run companies*. New York: Harper & Row, 1982
 311. Plsek P and Greenhalgh T. Complexity science: the challenge of complexity in health care. *British Medical Journal* 2001;**323**:625-628
 312. Pollitt C. Measuring performance: A new system for the National Health

- Service. *Policy and Politics* 1985;**13**(1):1-15
313. Pollitt C. Bringing consumers into performance measurement: concepts, consequences and constraints. *Policy and Politics* 1988;**16**(2):77-87
314. Pollitt C. The struggle for quality: the case of the NHS. *Policy and Politics* 1993;**21**(3):161-170
315. Pollitt C. Business approaches to quality improvement: why they are hard for the NHS to swallow. *Quality in Health Care* 1996;**5**(2):104-110
316. Pope C, van Toyen P and Baker R. Qualitative methods in research on healthcare quality. *Quality and Safety in Health Care* 2002;**11**:148-152
317. Popovich JA. Multidimensional performance measurement. *Journal of Nursing Care Quality* 1998;**12**(4):14-21
318. Popovich K and Popovich M. Use of Q-methodology for hospital strategic planning: a case study. *Journal of Healthcare Management*. 2000;**45**(6):405-414
319. Portelli B, Williams J and Collopy B. Using clinical indicators to change clinical practice. *Journal of Quality in Clinical Practice*. 1997;**17**:195-202
320. Powell RA and Single HM. Focus groups. *International Journal of Quality in Health Care* 1996;**8**(5):499-504
321. Power M. *The audit society: rituals of verification* Oxford: Oxford University Press, 1997
322. Power M. *The evolution of the audit society, its politics of control and the advent of CHI*. In Miles A, Hampton JR and Hurwitz B (eds.) *NICE, CHI and the NHS reforms: enabling excellence or imposing control?* London: Aesculapius

Press, 2000

323. Prasad RS. Development of the HIV / AIDS Q-sort instrument to measure physician attitudes. *Family Medicine*. 2001;**33**(10):772-778
324. Preston C and Baker R. Survey of the development of clinical governance in England and Wales. *Journal of Clinical Governance* 2000;**8**:118-123
325. Proctor S and Campbell J. A developmental performance framework for primary care. *International Journal of Health Care Quality Assurance* 1999;**12**(7):279-286
326. Quinn RE. *Beyond rational management*. San Francisco: Jossey-Bass, 1988
327. Quinn RE and Kimberly JR. *Paradox, planning and perseverance: guidelines for managerial practice* in JR Kimberly and RE Quinn (Eds) *Managing organizational translation* Homewood, IL: Dow Jones-Irwin, 1984
328. Quinn RE and Rohrbaugh J. A competing values approach to organizational effectiveness. *Public Productivity Review* 1981;**5**:122-140
329. Quinn RE and Rohrbaugh J. A spatial model of effectiveness criteria: toward a competing values approach to organisational analysis. *Management Science* 1983;**29**:363-377
330. Reymont R and Joreskog KG. *Applied factor analysis in the natural sciences*. New York: Cambridge University Press, 1993
331. Rhodes RAW. The new governance: governing without government. *Political Studies* 1996;**44**(4):652-667
332. Rhodes RAW. *Understanding governance: policy networks, governance,*

- reflexivity and accountability*. Buckingham: Open University Press, 1997
333. Rhodes RAW. *Foreword: governance and networks*, in Stoke G. (ed.) *The new management of British local governance*. Basingstoke: Macmillan, 1999
 334. Rigby KA, Palfreyman S and Michaels JA. Performance indicators from routine hospital data: death following aortic surgery as a potential measure of quality of care. *British Journal of Surgery* 2001;**88**:964-968
 335. Rissel C, Holt P and Ward J. Applying a health outcomes approach in a health service unit. *Australian Health Review* 1998;**12**(1):168-181
 336. Roberts A. Integrating strategy with performance measures. *Management Development Review* 1994;**7**(6):13-15
 337. Robinson R. Who's got the master card? *Health Service Journal* 2002;**112**(5824):22-24
 338. Robinson R. Change it or leave it. *Health Service Journal* 2003;**113**(5852):16
 339. Rogers CR. *Psychotherapy and personality change*. Chicago: University of Chicago Press, 1954
 340. Roland M, Campbell S and Wilkin D. Clinical governance: a convincing strategy for quality improvement? *Journal of Management in Medicine* 2001;**15**(3):188-201
 341. Rowan K and Black N. A bottom-up approach to performance indicators through clinician networks. *Health Care UK* 2001;(spring):42-46
 342. Rowland H. Janus: the two faces of the Commission for Health Improvement. *Clinical Governance: an International Journal* 2003;**8**(1):33-38

343. Runciman WB. Qualitative versus quantitative research – balancing cost, yield and feasibility. *Quality and Safety in Health Care* 2002;**11**:146-147
344. Russell E. The ethics of attribution: the case of health care outcome indicators. *Social Science and Medicine* 1998;**47**(9):1161-1169
345. Rust J and Golombok S. *Modern psychometrics: the science of psychological assessment*. London: Routledge, 1989
346. Salter B. Who rules? The new politics of medical regulation. *Social Science and Medicine* 2001;**52**:871-883
347. Scally G and Donaldson L. Clinical governance and the drive for quality improvement in the new NHS in England. *British Medical Journal* 1998;**317**:61-65
348. Schama S. *A history of Britain: at the edge of the world? 3000 BC – AD 1603*. London: BBC World-wide, 2000
349. Schaufeli WB, Leiter MP and Kalimo R. *The Maslach Burnout Inventory-General Survey: A self-report questionnaire to assess burnout at the workplace*. In MP Leiter, *Extending the burnout construct: Reflecting changing career paths*. Symposium, APA/NIOSH conference, Work, Stress and Health '95: Creating a healthier workplace. Washington, DC, 1995
350. Schein EH. *Organisational culture and leadership* San Francisco: Jossey-Bass, 1992
351. Schneider B. Organisational climates: an essay. *Personal Psychology* 1975;**28**:447-479

352. Schyve P. Models for relating performance measurement and accreditation.
International Journal of Health Planning and Management 1995;**10**(3):231-241
353. Scott JT, Mannion R, Davies HTO, Marshall MN. *Organisational culture and performance in the NHS: a review of the theory, instruments and evidence*. York: Centre for Health Economics, 2001
354. Scott JT, Mannion R, Davies HTO, Marshall MN. Does organisational culture influence health care performance? A review of the evidence. *Journal of Health Services Research and Policy* 2003a (in press).
355. Scott JT, Mannion R, Davies HTO, Marshall MN. The quantitative measurement of organisational culture in health care: what instruments are available? *Health Services Research* 2003b (in press).
356. Scriven M. Pros and cons about goal-free evaluation. *Evaluation Comment* 1972;**3**:1-7
357. Senn CY. Women's multiple perspectives and experiences with pornography. *Psychology of Women Quarterly*. 1993;**17**(3):319-341
358. Shapiro J. Seduced and betrayed by CHI. *Heath Service Journal* 2003;**113**(5839):18
359. Sharma S. Professional as agent: knowledge asymmetry in agency exchange. *Academy of Management Review* 1997;**22**(3):758-798
360. Shaw C. Health-care league tables in the United Kingdom. *Journal of Quality in Clinical Practice* 1997;**17**(4):215-219
361. Shaw CD. External quality mechanisms for health care: summary of the

- ExPeRT Project on visitatie, accreditation, EFQM and ISO assessment in European Union countries. *International Journal for Quality in Health Care*. 2000;**12**:169-175
362. Shaw I. *Evaluating public programmes: contexts and issues*. Aldershot: Ashgate Publishing, 2000
363. Sheldon T. Promoting health care quality: what role performance indicators? *Quality in Health Care* 1998;**7**(Suppl):s45-s50
364. Shiell A. Health outcomes are about choices and values: an economic perspective on the health outcomes movement. *Health Policy* 1997;**39**(1):5-15
365. Shirley J. Clinical governance in an independent hospital. *Clinician in Management* 2000;**9**:229-233
366. Shortell S, O'Brien JL, Carman JM, Foster RW, Hughes EFX, Boerstler H and O'Conner EJ. Assessing the impact of Continuous Quality Improvement/ Total Quality Management: Concept versus implementation. *Health Services Research* 1995;**30**(2):377-401
367. Shortell SM, Waters TM, Clarke KWB and Budetti PP. Physicians as double agents: maintaining trust in an era of multiple accountabilities. *Journal of the American Medical Association* 1998;**280**(12):1102-1108
368. Sieber SD. *Fatal remedies: the ironies of social intervention*. New York: Plenum Press, 1981
369. Singer SJ, Gaba DM, Geppert JJ et. al. The culture of safety: results of an organization-wide survey in 15 California hospitals. *Quality and Safety in Health*

Care 2003;**12**:112-118

370. Sitkin SB and Roth NL. Explaining the limited effectiveness of legalistic 'remedies' for trust / distrust. *Organizational Science* 1993;**4**(3):367-392
371. Smee C. The performance assessment framework: where did it come from and where is it going? *Health Care UK* 2000;(spring):3-8
372. Smith H and Frowen B. Assessing Success and Performance within the Public Sector. *The Mental Health Review* 1997;**2**(3):8-15
373. Smith J. The moderniser. *British Medical Journal* 2001;**323**:1150
374. Smith P. *Outcomes related performance indicators and organisational control in the public sector* In Holloway J. and Mallory G. Eds. *Performance measurement and evaluation* Sage 1995
375. Smith P. Performance indicators and outcome in the public sector. *Public Money and Management* 1995;**15**(4):13-16
376. Smith P. The unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration* 1995;**18**(2):277-310
377. Smith P. The use of performance indicators in the public sector. *Journal of the Royal Statistical Society A* 1990;**153**(1):53-72
378. Smith PC. Performance management in British health care: will it deliver? *Health Affairs* 2002;**21**(3):103-115
379. Smith PC, Stepan A, Valdmanis V and Verheyen P. Principal-agent problems in health care systems: an international perspective. *Health Policy* 1997;**41**:37-60
380. Snelling SJ. Women's perspectives on feminism: a Q-methodological study.

- Psychology of Women Quarterly*. 1999;**23**(2):247-266
381. Solberg LI, Mosser G and McDonald S. The three faces of performance measurement: Improvement, accountability and research. *Journal on Quality Improvement* 1997;**23**(3):136-147
 382. Springings N. Delivering public services under the new public management: the case of public housing. *Public Money and Management* 2002;**22**(4):11-17
 383. Spurgen P and De Luc C. *Development of clinical governance in PCTs in the West Midlands- lessons learned*. Birmingham: University of Birmingham Health Services Management Centre, 2002
 384. Stahr H. Developing a culture of quality within the United Kingdom healthcare system. *International Journal of Health Care Quality Assurance* 2001;**14**(4):174-180
 385. Stainton Rogers R. *Explaining health and illness: An exploration of diversity*. New York: Harvester Wheatsheaf, 1991
 386. Stainton Rogers R. *Q-methodology*. In Smith J, Harre R and Von Longenhove L. (eds.) *Rethinking methods in Psychology*. London: Sage, 1995 pp.178-192
 387. Stecher BM and Davies WA. *How to focus an evaluation*. London: Sage, 1987
 388. Stenner PHD, Dancey CP and Watts S. The understanding of their illness amongst people with irritable bowel syndrome: a Q methodological study. *Social Science and Medicine*. 2000;**51**:439-452
 389. Stephenson W. *The study of human behaviour: Q technique and its methodology*. Chicago: University of Chicago Press 1953

390. Stevens J. *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Lawrence Erlbaum, 1996
391. Stewart D and Shamdasani PN. *Focus Groups: theory and practice*. Newbury Park, CA: Sage, 1990
392. Stewart J and Walsh K. Performance measurement: when performance can never be finally defined. *Public Money and Management* 1994;**14**:45-49
393. Stordeur S, D'hoore W, Vandenberghe C. Leadership, organizational stress, and emotional exhaustion among hospital nursing staff. *Journal of Advanced Nursing* 2001;**35**(4):533-542
394. Streek W and Schmitter P. (eds.) *Private interest government*. London: Sage, 1985
395. Streiner GL and Norman DR. *Health measurement scales: a guide to their development and use*. Oxford: Oxford University Press, 1995 (2nd edition)
396. Stronach I. Shouting theatre in a crowded fire: 'educational effectiveness' as cultural performance. *Evaluation* 1999;**5**(2):173-193
397. Sweeney G, Sweeny K, Greco M and Stead J. *Developing clinical governance in primary care: Briefing paper 1: The experience of primary care clinical governance leads*. University of Exeter and Exeter and North Devon NHS Research and Development Support Unit, 2002
398. Sweeney G, Sweeny K, Greco M and Stead J. *Developing clinical governance in primary care: Briefing paper 2: Exploring the implementation and development of clinical governance within primary care in the SouthWest*. University of Exeter

- and Exeter and North Devon NHS Research and Development Support Unit, 2002b
399. Sweeney J, Heaton C. Interpretations and variations of ISO 9000 in acute health care. *International Journal for Quality in Health Care* 2000;**12**:203-209
400. Tabachnick BG and Fidell LS. *Using multivariate statistics*. London: Allyn and Bacon, 2001
401. Talbot C. Performing 'performance'- a comedy in five acts. *Public Money and Management* 2000;**20**(4):63-68
402. Tan JKH. The critical success factor approach to strategic alignment: seeking a trail from a health organization's goals to its management information infrastructure. *Health Services Management Research* 1999;**12**:246-257
403. Taormina RJ and Law CM. Approaches to preventing burnout: the effects of personal stress management and organizational socialization. *Journal of Nursing Management* 2000;**8**:89-99
404. Tarr JD. Performance measurements for a continuous improvement strategy. *Hospital Materiel Management and Quality* 1995;**18**(12):77-85
405. Taylor GB. What factors will facilitate the development of clinical governance in general practice? Results of a qualitative study. *Journal of Clinical Governance* 2000;**8**:181-185
406. Thomas WJ and Holloway J. Investigating early readmission as an indicator for quality of care studies. *Medical Care* 1991;**29**(4):377-394
407. Thompson G, Frances J, Levacic R and Mitchell J (eds.). *Markets, hierarchies and networks: the co-ordination of social life*. London: Sage, 1991

408. Thompson RG and Lally J. Performance management at the crossroads in the NHS: don't go into the red. *Quality in Health Care* 2000;**9**(1):201-202
409. Tinsley HEA and Tinsley DJ. Uses of factor analysis in counselling psychology research. *Journal of Counselling Psychology* 1987;**34**:414-424
410. Turpin RS, Darcy LA, Koss R, McMahon C, Meyne K, Morton D, Rodriguez J, Schmaltz S, Schyve P and Smith P. A model to assess the usefulness of performance indicators. *International Journal for Quality in Health Care* 1996;**8**(4)::321-329
411. Valenta AL and Wigger U. Q-methodology: Definition and application in health care informatics. *Journal of the American Medical Informatics Association*. 1999;**4**(6):501-510
412. Van der Bij JD and Vissers JMH. Monitoring health-care processes: a framework for performance indicators. *International Journal of Health Care Quality Assurance* 1999;**12**(5):214-221
413. Van Peursesem KA, Pratt MJ and Lawrence SR. Health management performance: a review of measures and indicators. *Accounting, Auditing & Accountability Journal* 1995;**8**(5):34-70
414. Vincent C. *Clinical risk management: enhancing patient safety* London: BMJ Books, 2001
415. Vincent C, Taylor-Adams S, Chapman EJ, Hewett D, Prior S, Strange P, Tizard A. How to investigate and analyse clinical incidents: Clinical Risk Unit and Association of Litigation and Risk Management protocol. *British Medical Journal*

2000;**320**:777-781

416. Wallace LM, Freeman T, Latham L, Walshe K and Sprurgeon P.
Organisational strategies for changing clinical practice: how trusts are meeting the challenges of clinical governance. *Quality and Safety in Health Care* 2001;**10**:76-82 (a)
417. Wallace L, Spurgeon P, Latham L, Freeman T, Walshe K. Clinical governance, organisational culture and change management in the new NHS. *Clinician in Management* 2001;**10**:23-31 (b)
418. Walsh MJ and Small N. Clinical governance in primary care: early impressions based on Bradford South and West Primary Care Group's experience. *British Journal of Clinical Governance* 2001;**6**(2):109-118
419. Walshe K. Improvement through inspection? The development of the new Commission for Health Improvement in England and Wales. *Quality and Safety in Health Care* 1999;**8**:191-196
420. Walshe K. Developing clinical governance: leadership, culture and change. *Journal of Clinical Governance* 2000;**8**:166-173
421. Walshe K. The rise of regulation in the NHS. *British Medical Journal* 2002;**324**:967-970
422. Walshe K and Dineen M. *Clinical Risk management: making a difference*. Birmingham: NHS Confederation, 1998
423. Walshe K and Freeman T. Effectiveness of quality improvement: learning from evaluations. *Quality and Safety in Health Care* 2002;**11**:85-87

424. Walshe K, Freeman T, Latham L, Wallace L and Spurgeon P. *Clinical governance: form policy to practice*. Birmingham: Health Services Management Centre, 2000
425. Webb C and Kevern J. Focus groups as a research method: a critique of some aspects of their use in nursing research. *Journal of Advanced Nursing* 2001;**33**(6):798-805
426. Wellard S. *Community Care* 26th October - 1 November 2000:10-11
427. White AW and Wager KA. The outcomes movement and the role of health information managers. *Topics in Health Information Management* 1998;**18**(4):1-12
428. Wilkin D, Coleman A, Dowling B and Smith K. *The national tracker survey of primary care groups and trusts 2001/2002: taking responsibility?* Manchester: NPCRDC, 2002
429. Wilkin D, Gillam S and Coleman A. (eds.) *The national tracker survey of primary care groups and trusts 2000/2001: modernising the NHS?* Manchester: NPCRDC, 2001
430. Wilkin D, Gillam S and Leese B. (eds.) *The national tracker survey of primary care groups and trusts 1999/2000: modernising the NHS?* Manchester: NPCRDC, 2000
431. Wilkin D, Hallam L, Doggett A. *Measures of need and outcome for primary health care*. Oxford: Oxford University Press, 1991
432. Williamson O. *The economic institutions of capitalism*. New York: Free Press, 1985

- 433. Wolff B, Knodel J and Sittitai W. *Focus groups and surveys as complementary research methods*. In Morgan D (ed.) *Successful focus groups: advancing the state of the art*. London: Sage, 1993: 118-136
- 434. Wright S. *Anthropology of organizations* London: Routledge, 1994
- 435. Wyszewianski L. The emphasis on measurement in quality assurance: reasons and implications. *Inquiry* 1988;**25**:424-436