MONITORING OF PATIENTS FOR THE DEVELOPMENT OF ADVERSE REACTIONS TO ANTIHYPERTENSIVE DRUGS IN GENERAL PRACTICE

by

SARAH ELIZABETH MCDOWELL

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

School of Clinical and Experimental Medicine
College of Medicine and Dental Sciences
University of Birmingham
May 2011

Abstract

Guidelines generally recommend the monitoring of serum electrolyte and creatinine concentrations in patients treated with antihypertensive therapy in order to detect potential adverse reactions to treatment. However, it is not well known to what extent these guidelines are followed in primary care.

I undertook a retrospective analysis of 74096 adult patients from the General Practice Research Database with newly diagnosed hypertension and prescribed a single antihypertensive agent. Baseline biochemical testing was undertaken in 31 094 patients (42%) and 37 365 (50%) patients had at least one biochemical monitoring test in the year after starting antihypertensive treatment. Monitoring was significantly more likely in patients treated with angiotensin-converting enzyme inhibitors than thiazide diuretics, older patients, and patients with diabetes mellitus. These patient factors were significantly associated with monitoring when multiple imputation was used to control for the potential bias introduced by missing data. In general, follow-up monitoring was infrequent, irregular, and did not change in response to events such as abnormal test results.

Patients who were monitored after the initiation of antihypertensive treatment were significantly more likely to be admitted to hospital and discontinue therapy, which is likely a result of reactive instead of planned monitoring. Using propensity score methods to control for confounding, I demonstrated a decreased risk of these same adverse outcomes in patients with baseline testing, which may be because these

patients were less likely to have any follow-up monitoring and not the protective effect of the baseline testing.

I described several barriers to biochemical monitoring including the lack of consensus in published guidelines, uncertain responsibility for monitoring, patient non-adherence, and absence of alerts or reminders to monitor. More work is needed to improve the primary evidence base for monitoring and to improve the guidelines on the nature and frequency of monitoring for adverse drug reactions, particularly in patients at greater risk of drug-induced harm.

For my family.

Acknowledgements

I am extremely grateful to my supervisor, Professor RE Ferner. I have learnt so much from Professor Ferner and his feedback, support, and guidance have been invaluable. I wish to thank and acknowledge Professor Jon Deeks for his readiness to share his expert knowledge of systematic reviews and statistics and for being so generous with his time.

I would also like to thank Dr Jamie Coleman, with whom a lot of the initial work for this thesis was carried out. His clinical knowledge, guidance, and energy have been of great help. Professor Stephen Evans from the London School of Hygiene and Tropical Medicine has provided me with significant guidance and help with statistical concerns and queries, for which I am most grateful. He introduced me to the concept of immortal time bias and underscored the need to adjust for this type of bias, which is often ignored in various types of research. I would also like to acknowledge Dr Paramjit Gill for his help with the identification and further refinement of appropriate clinical codes.

I am very grateful to my former colleagues at the West Midland Centre for Adverse Drug Reactions for their assistance and support. Mr Christopher Anton assisted me with database manipulation and design and Dr Anthony Cox provided valuable comments and discussions.

The study protocol was approved by the Independent Scientific Advisory Committee for Medicines and Healthcare products Regulatory Agency database (MHRA)

research. This study is based in part on data from the Full Feature General Practice Research Database obtained under licence from the MHRA. However, the interpretation and conclusions contained in this thesis are mine alone.

Table of contents

List of figures

List of tables

List of appendices

List of abbreviations

ACE	Angiotensin-converting enzyme
ADR	Adverse drug reaction
AT-II	Angiotensin-II
BMI	Body mass index
BNF	British National Formulary
CI	Confidence interval
GoF	Goodness-of-fit
GP	General Practitioner
GPRD	General Practice Research Database
HIT	Health information technology
HMG Co-A	3-Hydroxy-3-methylglutaryl-coenzyme A
HMO	Health maintenance organization
HR	Hazard ratio
ISAC	Independent Scientific Advisory Committee
IQR	Interquartile range
MAR	Missing at random
MCAR	Missing completely at random
MCMC	Markov Chain Monte Carlo
MHRA	Medicines and Healthcare products Regulatory Agency
MICE	Multiple imputation by chained equations
MNAR	Missing not at random
MPR	Medication possession ratio
MRC	Medical Research Council
NHS	National Health Service
NSAID	Non-steroidal anti-inflammatory drug
OR	Odds ratio
OXMIS	Oxford Medical Information System
RCT	Randomized controlled trial
ROC	Receiver operator characteristic
RR	Relative risk
SD	Standard deviation
SES	Socio-economic status
SPC	Summary of Product Characteristics
SSRI	Selective serotonin re-uptake inhibitors
UK	United Kingdom
UTS	Up-to-standard
WHO	World Health Organization

Chapter 1

INTRODUCTION

This chapter provides a foundation for the work presented in the other chapters in this thesis. I describe the nature of adverse drug reactions in general terms and in relation to the use of antihypertensive therapy. I also introduce the concept of monitoring to reduce the risk of harm due to adverse reactions to therapy. I present a summary of the recommendations from published guidance on the monitoring of patients treated with antihypertensive therapy for adverse drug reactions. Finally, I describe the overall aims and objectives of the thesis.

1.1 Adverse drug reactions

Definition

An adverse drug reaction (ADR) has been defined by the World Health Organization (WHO) as: 'any response to a drug which is noxious, unintended and occurs at doses used for prophylaxis, diagnosis, or therapy.' ADRs can be divided into two types: those reactions that are predictable and dose dependent (Type A) and those which are unpredictable and unusual reactions (Type B). Type B reactions are less common but can be more serious than type A reactions. This classification is simple and has been

widely used but has also been criticized as it is sometimes difficult to assign an ADR to only one type.

More recently, a three-dimensional classification system for the analysis of ADRs has been suggested based on dose relatedness, timing, and patient susceptibility (DoTS) (Aronson & Ferner, 2003). When such a classification system is used, the probability of an ADR may vary depending on the time after drug administration, the dose of the drug, and various patient susceptibility factors such as age, sex, genetic factors, and exogenous factors.

Incidence of ADRs

Determining the incidence of ADRs is challenging due to several factors including: (1) different definitions of an ADR; (2) different methods used to identify, evaluate and document ADRs; and (3) different population groups. Several large systematic reviews of ADRs have demonstrated that as few as 0.16% to as many as 15.7% of hospital admissions are due to ADRs (Lazarou *et al.*, 1998; Impicciatore *et al.*, 2001; Beijer & de Blaey, 2002; Wiffen *et al.*, 2002; Davies *et al.*, 2007; Kongkaew *et al.*, 2008). The incidence of fatal ADRs in patients admitted to hospital has been reported as ranging from 0.05% to 0.44% (Pouyanne *et al.*, 2000; Zoppi *et al.*, 2000; Ebbesen *et al.*, 2001; Juntti-Patinen & Neuvonen, 2002; Mjorndal *et al.*, 2002; Schneeweiss *et al.*, 2002). Although the range in the estimated incidence of ADRs does vary widely, adverse reactions to therapeutic treatment remain a cause of significant morbidity and mortality.

The incidence of ADRs in the community is less well known, with fewer well-designed studies. A small study of a single general practice demonstrated that 1.7% of routine consultations or extra appointments were due to an ADR, the majority being due to three common groups of drugs: antidepressants, antibiotics, and non-steroidal anti-inflammatory drugs (Millar, 2001). Gurwitz and colleagues (2003) identified an incidence of 50.1 adverse drug events per 1000 person-years in a group of adults in ambulatory care, with almost a third (28%) deemed to have been 'preventable'. A study published in the same year reported that 25% of patients in the community had an ADR (Gandhi *et al.*, 2003). The incidence of fatal ADRs in a large population study in Sweden was found to be 3%, which is higher than reported in hospital (Wester *et al.*, 2008).

ADRs in the UK

The largest prospective study of ADRs to be carried out in the UK determined that 6.5% of hospital admissions were related to ADRs, with the ADR directly leading to the admission in 80% of cases (Pirmohamed *et al.*, 2004). Almost three-quarters (72%) of cases were classified as being preventable. The incidence of fatal ADRs in hospitalized patients was 0.15%, similar to other reports. The impact of ADRs in the NHS in England has been estimated to be 4 out of 100 hospital bed-days, at a cost of about £380 million a year (Wiffen *et al.*, 2002).

Risk factors for ADRs

ADRs tend to be more common in older adults, which may reflect the higher prevalence of long-term disease and poly-pharmacy in this age group. Young children are also at risk of ADRs, mainly because doses need to be calculated individually

based on the patient's age, weight or body surface area, and clinical condition (Wong *et al.*, 2004). In general, the risk of ADRs is greater in women (1.5–1.7 greater risk) (Rademaker, 2001). ADRs may also be more common in patients with concomitant renal, hepatic, and cardiac disease.

1.2 Adverse drug reactions associated with antihypertensive drugs

Antihypertensive drugs are used to treat patients with sustained elevation of systolic or diastolic blood pressure. In the early 1990s there was a significant increase in the use of newer agents such as angiotensin-converting enzyme (ACE) inhibitors and angiotensin-II (AT-II) receptor antagonists (Psaty *et al.*, 1995). However, thiazide diuretics—usually considered to be more traditional agents—still remain the most frequently used drug class (one in three patients) for the treatment of patients with newly diagnosed hypertension (Walley *et al.*, 2003). Indeed, thiazide diuretics, along with calcium (Ca)-channel blockers are recommended as initial therapy in patients aged 55 and over or in non-white patients of any age in the UK. In hypertensive patients younger than 55, ACE inhibitors should be used as first-line therapy (National Collaborating Centre for Chronic Conditions, 2006).

Antihypertensive drugs are generally well tolerated but are known to cause a range of ADRs, particularly electrolyte disturbances such as hyperkalaemia, hypokalaemia, and hyponatraemia (Olsen *et al.*, 1999; Liamis *et al.*, 2008a).

1.2.1 Hyperkalaemia and antihypertensive drugs

Hyperkalaemia is a condition of excess potassium, which can lead to abnormal heart rhythm, muscle spasm, cardiac arrest, and potentially death. Moderate hyperkalaemia is defined as a serum potassium concentration greater than 5.5 mmol/l; severe hyperkalaemia as a concentration greater than 6.0 mmol/l. The incidence of hyperkalaemia in hospitalized patients varies from 1–10% depending on the definition used (Liamis *et al.*, 2008a).

ACE inhibitors are responsible for 10–38% of cases of hyperkalaemia in hospitalized patients (Liamis *et al.*, 2008a) and often cause hyperkalaemia by inducing a state of hypoaldosteronism or by impairing renal potassium excretion (Rimmer *et al.*, 1987). The combination of ACE inhibitors with other potential potassium-altering medications such as potassium-sparing diuretics can also increase the risk of hyperkalaemia (Chiu *et al.*, 1997; Saito *et al.*, 2005). Hyperkalaemia is more likely to occur in older patients, patients with impaired renal function or hypoaldosteronism, and those with diabetes mellitus (Ramadan *et al.*, 2005; Indermitte *et al.*, 2007).

Potassium-sparing diuretics (e.g. amiloride, triamterene, and spironolactone) may precipitate hyperkalaemia by diminishing potassium secretion and have been associated with 4–19% of cases of moderate to severe hyperkalaemia (Perazella, 2000). Non-selective beta-adrenoceptor blocking drugs (beta-blockers) have also been associated with hyperkalaemia, although the condition is rarely severe. Beta-blockers may lead to hyperkalaemia through the suppression of catecholamine-stimulated renin release, which decreases aldosterone synthesis; and by decreasing cellular uptake of potassium (Perazella, 2000).

1.2.2 Hypokalaemia and antihypertensive drugs

Hypokalaemia refers to an abnormally low serum potassium concentration and is generally defined as a serum concentration less than 3.5 mmol/l. The most common drug-related cause is diuretic therapy and the degree of hypokalaemia is greater with higher doses of diuretics and increased dietary sodium intake (Gennari, 1998; Liamis *et al.*, 2008a).

The incidence and severity of hypokalaemia with low-dose thiazide diuretic therapy is relatively low. One percent of patients in the large Systolic Hypertension in the Elderly study had a serum potassium concentration less than 3.2 mmol/l (SHEP Cooperative Research Group, 1991). Although rare, hypokalaemia can lead to potentially dangerous cardiac arrhythmias and may also be associated with an increased risk of cardiovascular events (Freis, 1995; Moser, 1998).

The relationship between diuretics and hypokalaemia is less well understood. The risk is increased in patients with liver cirrhosis and in patients with severe cardiac failure that is complicated by secondary hyperaldosteronism (Aronson, 2006). The risk of hypokalaemia is greater in older patients, although this may reflect the increased use of thiazide diuretics in this population group (Zuccalà *et al.*, 2000). Diuretic-induced hypokalaemia is usually seen in the first few weeks following initiation of treatment (Siegel *et al.*, 1992; Miltiadous *et al.*, 2003).

1.2.3 Hyponatraemia and antihypertensive drugs

Hyponatraemia is an ADR commonly associated with hospitalization and is identified by an abnormally low serum sodium concentration. Definitions vary with some researchers using a definition of less than 135 mmol/l, while others have used less than or equal to 130 mmol/l. Early symptoms include nausea, headache, muscular weakness and lethargy, and if untreated the condition can progress to seizures, neurological damage, and death (Clayton *et al.*, 2006a).

Diuretics are one of the most common drug classes associated with hyponatraemia and the majority of cases are caused by thiazide diuretic or thiazide diuretic-like agents (Spital, 1999; Chow *et al.*, 2003; Clayton *et al.*, 2006a; Liamis *et al.*, 2008b). Diuretic-induced hyponatraemia is more likely to occur in older women (Sharabi *et al.*, 2002), although this relationship may only reflect the high levels of diuretic prescription in this demographic group. Cases of thiazide diuretic-induced hyponatraemia have also been shown to develop most often within two weeks of starting treatment (Spital, 1999; Aronson, 2006).

1.3 Monitoring patients for adverse drug reactions

Monitoring is a process of checking a system that changes with time, in order to guide changes to the system that will maintain or improve it (Coleman *et al.*, 2006). Monitoring of drug treatment can have several effects including better selection of drug therapy, better titration of treatment, and improved adherence (Glasziou *et al.*, 2005). Monitoring can also—perhaps most importantly—identify potential adverse reactions to treatment. Failure to monitor renal function and electrolytes has been

shown to be responsible for 26% of preventable drug-related hospital admissions (Howard *et al.*, 2003) and 36% of preventable adverse drug events in ambulatory care (Thomsen *et al.*, 2007).

Ideally, a test to monitor for an ADR should be safe, simple, accurate, and acceptable to both patients and health care professionals (Pirmohamed & Ferner, 2003; Ferner *et al.*, 2005). The test should also be cost proportionate, in that the cost of testing should not be greater than the cost savings associated with reducing the health burden of adverse reactions to treatment.

If possible, monitoring instructions should provide details on:

1. The purpose of the monitoring;
2. The types of tests to be used;
3. When to start monitoring;
4. A satisfactory frequency of monitoring (although this may not necessarily be constant);
5. An acceptable range of values;
6. Actions to be taken should the test identify a value outside of the stated acceptable range;
7. When, if appropriate, it is safe for monitoring to cease.

Specific guidance on monitoring for ADRs available to health care professionals is often limited. Guidelines exist from a wide variety of sources, such as publications from government organizations, professional societies (e.g. British Hypertension

Society), or independent researchers or research groups, but often vary in their quality and level of detail on monitoring. Often, published guidelines do not provide clear recommendations for the number of monitoring tests or guidance on the interpretation of subsequent monitoring. Monitoring guidance sent to general practitioners (GPs) when therapeutic treatment is initiated in hospital is often incomplete (Corry *et al.*, 2000).

Drug manufacturers also provide some information on monitoring for ADRs in the Summary of Product Characteristics (SPC) for each drug. A SPC is presented in a standardized format and serves as both a regulatory document and as a primary source of information for health professionals (Waller & Evans, 2003). However, the latter goal is often disadvantaged by the former, as the limitations placed on presentation and the lack of flexibility imposed by the standardized format limit how information is presented (Waller & Evans, 2003). Certainly, the information on monitoring for ADRs in SPCs has been described as vague and inadequate in that it does not provide enough information to carry out any monitoring or act on the results (Ferner *et al.*, 2005).

1.4 Monitoring guidelines for adverse reactions to drugs used in the treatment of hypertension

Hypertension is one of the commonest conditions managed within primary care (National Collaborating Centre for Chronic Conditions, 2006). Drugs used in the treatment of hypertension are generally well tolerated but are known to cause a range of ADRs (section 1.2). Specifically, renal failure can occur with ACE inhibitors, AT-II receptor antagonists and diuretics. Hyperkalaemia is also a recognized adverse

effect of ACE inhibitors, AT-II receptor antagonists and potassium-sparing diuretics and hypokalaemia with thiazide and loop diuretics, which can also cause hyponatraemia. Biochemical monitoring after the initiation of antihypertensive therapy can consequently identify changes related to ADRs before they have caused serious or permanent effects, and so avert them. Monitoring for ADRs that may not necessarily manifest symptoms, such as in most cases of hypokalaemia (Gennari, 1998), is of significant importance.

In order to understand the nature of the published guidance on monitoring, I carried out a purposeful, although not exhaustive, search to identify (1) guidelines written with the purpose of describing the monitoring of patients treated with antihypertensive therapy or (2) guidelines on the treatment of patients with hypertension. In January 2010 I searched Medline using a combination of text and MeSH terms for hypertension and drugs used in the treatment of hypertension, limiting the results to the publication type 'practice guideline'. I searched the National Guideline Clearinghouse (www.guideline.gov) database using the search term hypertension. I also searched the electronic Medicines Compendium (http://www.medicines.org.uk/emc/), to identify monitoring recommendations listed in SPCs for drugs licensed in the UK and used in the treatment of hypertension. I used the following search terms to identify the recommendations: monitoring, renal function, test or testing, creatinine, urea, potassium, sodium, or urea. I did not search the SPCs of combination products, and all of the SPCs for a given drug were reviewed when the drug was manufactured by multiple companies.

I identified 23 publications that contained guidance on the monitoring of patients treated with antihypertensive therapy, of which three were written especially for patients with chronic kidney disease (Appendix 1). I also identified five well-cited guidelines on the use of ACE inhibitors, AT-II receptor antagonists or spironolactone in the treatment of heart failure. All of the guidelines provided information on what serum concentration to monitor. I reviewed the SPCs for 58 different antihypertensive drugs (Appendix 2).

The recommendations on the frequency of monitoring varied in their level of detail. Some guidelines described monitoring as a 'routine investigation' or that laboratory tests should be 'measured annually'. Other guidelines were far more detailed with recommendations for biochemical testing prior to the initiation of therapy, specific details for the frequency of follow-up monitoring, and actions to be taken should the laboratory tests be outside a certain range of concentrations (French Haute Autorité de Santé, 2005; Northern Ireland Department of Health Social Services and Public Safety, 2007; Smellie *et al.*, 2007). The most detailed recommendations for monitoring were provided in the guidelines for the treatment of hypertension in patients with chronic kidney disease (National Kidney Foundation, 2002; Department of Veterans Affairs/ Department of Defense, 2007; National Institute for Health and Clinical Excellence, 2008). In addition to data on the frequency of monitoring and when to stop treatment, these guidelines also provided information on how to tier the frequency of monitoring based on the initial baseline serum concentration.

Most of the guidelines did not explicitly reference the primary evidence supporting the recommendations, as has also been shown previously in an analysis of renal

function monitoring guidelines in patients treated with ACE inhibitors (Coleman *et al.*, 2008). Indeed, monitoring guidelines for antihypertensive therapy have been previously criticized: Sica (2006) described the guidelines for the monitoring of serum potassium concentration in patients treated with antihypertensive drugs as 'at best makeshift and often drawn from the know-how of the treating physician'.

The monitoring recommendations made by the drug manufacturers in the SPCs were, in general, rather vague and lacked detail. This is similar to the conclusions made by Ferner and colleagues (2005), who concluded that the monitoring recommendations for haematological reactions for non-oncological drugs were inadequate. Although some SPCs specifically recommended monitoring in the first few weeks of treatment, the majority were limited to statements describing 'regular' or 'routine' monitoring. Some SPCs did suggest more vigilant monitoring in patients at increased risk of drug-induced harm (e.g. patients with renal impairment, older patients, or patients with diabetes mellitus). There was some additional detail on which drugs required additional monitoring when used concomitantly with antihypertensive therapy due to the increased risk of biochemical ADRs or renal impairment. No monitoring recommendations were provided for alpha-blockers, beta-blockers, or calcium-channel blockers.

The nature and detail of published guidelines and recommendations for monitoring differs, which may be due to the lack of primary evidence. However, it is clear that a general consensus exists: patients with newly diagnosed hypertension should have baseline biochemical tests of renal function and electrolyte concentrations before

treatment, monitored again after starting treatment, and at intervals following any dose changes.

1.5 Aims and objectives of the thesis

Prior studies have demonstrated that monitoring of patients treated with antihypertensive drugs for electrolyte disturbances is often not undertaken. Few published studies have put their results into context with existing guidelines on monitoring and even fewer have examined the relationships between monitoring and patient outcomes.

The overall aim of this study is to understand the nature of monitoring for ADRs in a population of patients newly diagnosed with hypertension and newly treated with antihypertensive drugs using the General Practice Research Database (GPRD).

Specifically, the objectives of my research are:

1. To observe and clarify the patterns of monitoring for ADRs in patients treated with antihypertensive drugs;
2. To determine relationships between the outcomes for patients, the patterns of monitoring for ADRs, and other factors;
3. To compare the observed patterns of monitoring of patients after the initiation of antihypertensive drugs for ADRs against published recommendations for monitoring.

1.6 Outline of the thesis

The use of the GPRD—a very large and well-validated database—will provide a large number of patients and clinical data. This analysis will present information on the monitoring of patients for ADRs in general practice. Patients who are newly treated with antihypertensive drugs may be at greater risk of harm and ADRs, and will be specifically examined. At present, this area is under-researched and the analysis will provide important data with which to inform strategy for monitoring the safe use of drugs in representative populations. The chapters of the thesis are as follows:

- Chapter 2 presents the results of a systematic review of studies examining biochemical monitoring in patients treated with antihypertensive drugs;
- Chapter 3 introduces the GPRD and describes the development of the study cohort;
- Chapter 4 describes the statistical methods used to analyze the nature of biochemical laboratory monitoring, assessing the frequency, regularity, and responsiveness of biochemical laboratory tests;
- Chapter 5 presents the results of the analyses described in Chapter 4;
- Chapter 6 examines the nature of biochemical monitoring and the frequency of hyponatraemia in a sub-group of thiazide diuretic-treated patients;
- Chapter 7 describes the method of multiple imputation as a tool for dealing with missing data and compares the results obtained using multiple imputed data with those from a complete case analysis;
- Chapter 8 introduces propensity score methods for controlling for confounding in observational research, which are used in addition to traditional methods to

examine the relationship between baseline biochemical testing and adverse patient outcomes;

- Chapter 9 describes the final conclusions and details future work.

Chapter 2

BIOCHEMICAL LABORATORY MONITORING OF PATIENTS TREATED WITH ANTIHYPERTENSIVE DRUGS: A SYSTEMATIC REVIEW

This chapter presents the results of a systematic review of the nature of biochemical laboratory monitoring in patients treated with antihypertensive therapy. I demonstrate that previous research, primarily undertaken in the United States, has shown a lack of biochemical monitoring in primary care. I discuss the strengths and weaknesses of prior research and demonstrate the significant lack of research on the relationship between monitoring and adverse patient outcomes.

2.1 Background and research questions

Biochemical monitoring of patients treated with antihypertensive therapy is recommended by many guidelines (Appendix 1). Tests of serum creatinine and electrolyte concentrations can identify adverse reactions to antihypertensive treatment such as renal failure and hyponatraemia. However, it is not well known to what extent these guidelines are followed in primary care. It is also unclear whether monitoring

reduces the risk of adverse patient outcomes. Therefore I wished to identify and summarize the literature investigating the nature of biochemical monitoring in adults treated in primary care with antihypertensive drugs. This was undertaken by addressing five research questions:

1. What is the proportion of patients with biochemical testing prior to the initiation of antihypertensive therapy?
2. What is the proportion of patients with biochemical monitoring after initiation of antihypertensive therapy?
3. What are the patient characteristics associated with biochemical monitoring?
4. What is the frequency of biochemical monitoring after the initiation of antihypertensive therapy?
5. What is the relationship, if any, between biochemical monitoring and adverse patient outcomes?

2.2 Selection criteria

Types of studies	Randomized controlled trials or quasi-randomized controlled trials investigating the impact of various interventions on monitoring were included. The results from the control arm of a randomized trial were used when comparisons were made with other studies. Retrospective and prospective cohort studies, cross-sectional studies, and audits of current clinical practice were also included. Clinical trials of drug therapy were excluded because they do not accurately reflect monitoring in a normal treatment setting. Case reports and case series were also excluded.
Types of patients	Male and female adults with hypertension treated with antihypertensive therapy outside of the hospital or clinical trial setting were included. No upper age limit was applied.

Types of biochemical monitoring	The biochemical monitoring tests identified for inclusion were tests of serum creatinine, urea, sodium, or potassium concentration.
Types of outcome measures	The primary outcome measure was • The proportion of patients with any biochemical testing prior to the start of treatment or follow-up monitoring. The secondary outcome measures were • The frequency of tests (number of tests over a period of time); • Patient factors associated with biochemical testing at baseline or follow-up monitoring; • Any additional information reported on baseline testing or biochemical monitoring.

2.3 Search strategy for the identification of studies

2.3.1 Electronic search

Two electronic databases—Medline and Embase—were searched using the OVID interface to identify studies for potential inclusion. I used a combination of text words and controlled vocabulary search terms for Medline (MeSH) and Embase (EMTREE) (Appendix 3). The development of the search strategy was an iterative process where an initial analysis of the relevant search terms was carried out. This was done in order to identify superfluous or ineffective search terms and to create a search strategy that was both sensitive and specific. The strategy was further tested by checking whether it could identify relevant studies that had been previously identified. The search was carried out with no language restrictions.

Once duplicates were removed from the combined databases, studies were selected for inclusion based on the study title and abstract. When studies appeared to meet the

inclusion criteria or where a decision could not be made based solely on title or abstract, full-text copies were obtained.

Citation searching

Articles that cited or were cited by the included studies were also screened in order to identify any further relevant studies. Additionally, reference lists from important reviews were searched and personal files were examined in order to identify further studies.

Identifying any unpublished literature

Google and Google Scholar were also searched in order to identify studies that had not been published or had been published in journals not indexed by Medline or Embase using a combination of search words:

> (antihypertensive OR alpha adrenoceptor blocking OR alpha blockers OR adrenoceptor blocking OR alpha blockers OR aldosterone antagonists OR diuretics OR ACE inhibitors OR angiotensin converting enzyme inhibitors OR angiotensin receptor blockers OR angiotensin II receptor antagonists OR calcium channel blockers) AND (monitoring OR serum potassium OR serum sodium OR serum creatinine OR serum urea OR potassium concentration OR sodium concentration OR creatinine concentration OR urea concentration)

2.3.2 Methods of the review

Data extraction was carried out using a standardized data extraction form (Appendix 4). A narrative summary was used to describe each study's methods and key findings. The primary outcome measure—the proportion of patients with any biochemical testing or follow-up monitoring—was compared between the included studies. If data were presented in a randomized trial examining the effect of various interventions on monitoring, only data from the control group were extracted.

I set out a list of quality and methodological items to assess the studies identified by the review. Using the Newcastle-Ottawa Quality Assessment Scale (NOS) (Wells *et al.*, 2010) and the Quality Assessment of Diagnostic Accuracy Studies tool (Whiting *et al.*, 2003) as guides, a list of five quality indicators was developed in order to critically appraise the included studies. Seven additional methodological indicators applicable to the understanding of the nature of biochemical monitoring were also included in the appraisal process.

2.4 Results

In January 2010 I searched Medline (from 1948) and Embase (from 1980), without any language restrictions, using a combination of medical subject headings and text words. The initial search strategy retrieved 1223 studies. 91 studies were selected based on their title and abstract and full-text copies were retrieved. Four studies were identified through hand-searching (Hurley *et al.*, 2005; Raebel *et al.*, 2005; Feldstein *et al.*, 2006; Palen *et al.*, 2006; Sauer *et al.*, 2006; Matheny *et al.*, 2008).

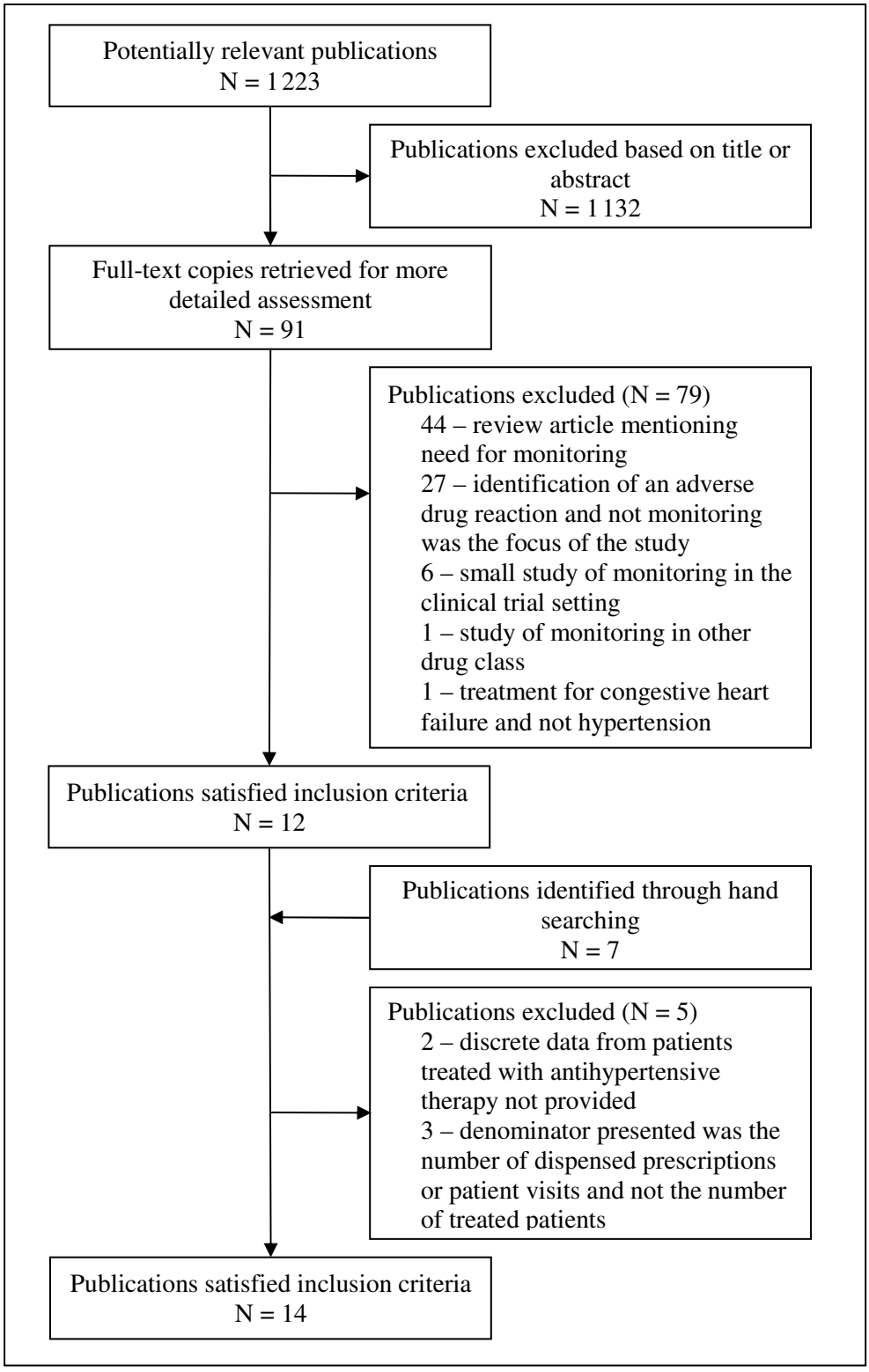


Figure 2.1 – Flow chart of selection of studies into systematic review

Two studies that examined interventions for increasing laboratory monitoring were excluded as they did not provide discrete data on patients treated solely with

antihypertensive therapy (Steele *et al.*, 2005; Feldstein *et al.*, 2006). Three studies were also later excluded as they presented only data on the proportion of dispensed prescriptions or patient visits with biochemical monitoring, and not the proportion of patients (Raebel *et al.*, 2005; Palen *et al.*, 2006; Matheny *et al.*, 2008). Fourteen studies were included in the final review (Figure 2.1). A detailed summary of the studies selected for inclusion is provided in Appendix 5 and each study is considered individually below. The excluded studies are described in Appendix 6.

2.4.1 Narrative review of included studies

An early cross-sectional study was carried out in a UK general practice of 330 patients treated with diuretics (Rhodes, 1992). Almost one quarter (23%) of patients had no record of any monitoring of urea and electrolyte concentrations. 36 patients (11%) had a record of urea and electrolyte monitoring prior to the initiation of diuretic treatment but not subsequently.

A postal questionnaire of 277 British GPs was carried out to determine the proportion of GPs who regularly monitored renal function in patients treated with ACE inhibitors. 85% of the GPs reported that they checked renal function prior to initiation of treatment, and fewer GPs (34%) monitored after treatment (Kalra *et al.*, 1999). An audit of the laboratory records from 122 patients in one GP practice by the same authors revealed that GPs tended to overestimate how often they carried out any monitoring: only 45% of patients had their renal function monitored before ACE inhibitor treatment and 29% had any monitoring subsequent to treatment initiation.

A large study of an Israeli computerized patient record system identified approximately 35 000 patients who were treated with a diuretic (Hoch *et al.*, 2003). Almost 22% of patients did not have a record of a serum potassium concentration test within one year of the start of treatment. This value decreased significantly to 18% after the initiation of a computer alert sent to physicians indicating that a patient had not had a monitoring test.

The large retrospective analysis of data from a US Health Maintenance Organization (HMO) by Hurley and colleagues (2005) examined the rates of missed laboratory tests in patients treated with ACE inhibitors and diuretics from 1999 to 2001. Approximately a third of patients (35%) did not have any creatinine or potassium monitoring within one year of starting treatment with an ACE inhibitor. Similar rates were observed with patients treated with diuretics. The rates of monitoring were shown to improve over the three-year period.

A cross-sectional analysis of patients from the United States aged over 65 years examined the proportion of patients treated with cardiovascular medications who did not have any baseline testing (defined as a creatinine or potassium test in the 180 days before and 14 days after the start of treatment) (Simon *et al.*, 2005). Approximately one third of patients had no evidence of any baseline testing. In patients treated with diuretics, biochemical testing was less likely to occur in women and patients with fewer co-morbidities. No sex difference was observed in patients treated with ACE inhibitors or AT-II receptor antagonists.

Clayton and colleagues (2006b) undertook a large, retrospective analysis of electronic prescribing and laboratory records of patients who were treated with at least one thiazide diuretic from six UK GP practices. 17% of patients had at least a sodium or potassium concentration test in the two years before initiation of thiazide diuretic treatment. Almost a third of patients had at least one follow-up monitoring test of sodium or potassium within two years of the start of treatment. Male sex, increasing age, and more concomitant prescriptions were shown to be statistically significantly associated with electrolyte monitoring in the two years after the initiation of antihypertensive treatment.

An analysis of records from a large veterans database in the United States reported that 81% of patients had at least one baseline potassium or creatinine test prior to the initiation of ACE inhibitor or AT-II receptor antagonist treatment (Sauer *et al.*, 2006). Half of the cohort had both baseline testing and follow-up monitoring within 12 weeks. However, over two thirds of the cohort (67%) did not have a baseline test and follow-up monitoring within four weeks of treatment. Patients with fewer outpatient encounters (OR 1.9, 95% CI 1.5–2.3), and those travelling longer distances to receive treatment (OR 1.19, 95% CI 1.03–1.4) were significantly more likely not to have follow-up monitoring.

Lafata and colleagues (2007) undertook a cluster-randomized study in primary care practices to examine the effect of an academic detailing intervention with performance feedback on baseline testing in patients newly treated with antihypertensive therapy and continuing users of the treatment. The intervention involved two face-to-face visits where the primary care providers received feedback

on their current rates of laboratory monitoring, research evidence supporting the need for monitoring, and strategies to improve monitoring within their practice. Pre-intervention, 59% of patients in the control cohort treated with an ACE inhibitor or AT-II receptor antagonist had both a baseline serum creatinine and potassium test. 55% of patients treated with a thiazide diuretic had any baseline potassium testing.

The very large study by McAlister and colleagues (2007) examined administrative drug databases from one Canadian province for 164 413 patients aged 66 years and older who were newly treated with antihypertensive monotherapy. Overall, 41% of patients had at least one laboratory test in the six months before initiation of treatment and 49% of patients had at least one laboratory test during one year of follow-up. Patients prescribed thiazide diuretics were more likely to have their serum electrolytes measured at least once during follow-up, compared with newer agents (ACE inhibitors, AT-II receptor antagonists, and Ca-channel blockers). Thiazide diuretic-treated patients were less likely to have any renal function monitoring than patients treated with newer agents.

Raebel and colleagues (2007a) undertook a retrospective analysis of data from 10 HMO databases of patients treated with spironolactone. Almost three quarters of patients had both serum creatinine and potassium tests within one year of treatment. Male sex, increasing age, and diabetes mellitus were some of the patient covariates associated with monitoring.

Another large retrospective analysis of over 50 000 patients treated with ACE inhibitors or AT-II receptor antagonists was carried out using data from 10 HMOs

(Raebel *et al.*, 2007b). 64% of the cohort had at least one creatinine and one potassium concentration test within one year of continuing therapy. Patient factors such as increasing age, male sex and co-morbidities such as diabetes mellitus and chronic renal failure were significantly associated with laboratory monitoring.

Besançon and colleagues (2008) undertook a review of computerized patient records from a French health insurance database for 3620 patients treated with a spironolactone-ACE inhibitor combination. In the six months prior to the initiation of treatment, 1083 patients (30%) had evidence of at least one record of both serum potassium and creatinine measurements.

A retrospective study analyzed of records from older adults in a French national insurance database who were treated with diuretics (Gérardin-Marais *et al.*, 2008). 70% of the population had at least one serum chemistry monitoring test within one year of treatment. Older age, female sex and serious disease were significantly associated with monitoring.

The most recent study was an analysis of data from diabetic patients treated with an ACE inhibitor, an angiotensin-II receptor antagonist or spironolactone (Raebel *et al.*, 2010). 71% of the cohort had at least one serum potassium concentration measured during treatment. Patients who were monitored were more likely to be older, female, and have more chronic disease. Serum potassium monitoring was associated with a decreased risk (adjusted RR 0.50, 95% CI 0.37–0.66) of hyperkalaemia and hyperkalaemia-associated adverse events such as emergency department visits, hospitalization and death. The risk of adverse events further decreased with

monitoring in a sub-group of patients with chronic kidney disease (adjusted RR 0.19, 95% CI 0.11–0.36).

2.4.2 Summary of studies

The UK, France, and the USA provided the majority of the studies and all but two of the studies were cross-sectional or retrospective analyses of monitoring (Hoch *et al.*, 2003; Lafata *et al.*, 2007). All of the studies used databases of electronic prescription records to identify the patients treated with antihypertensive therapy. The majority of studies (79%) used electronic records to identify laboratory tests, with only one study undertaking an assessment of the ability of the administrative data to identify monitoring (Raebel *et al.*, 2007b).

Most studies (71%) were undertaken using insurance databases to identify a range of patients typically treated with antihypertensive therapy in primary care. Some studies focused on specific sub-groups of patients, such as the elderly (Simon *et al.*, 2005; Gérardin-Marais *et al.*, 2008) or those with diabetes (Raebel *et al.*, 2010). Three studies were carried out in patients newly treated with antihypertensive therapy (Lafata *et al.*, 2007; McAlister *et al.*, 2007; Raebel *et al.*, 2010). The majority of studies focused on follow-up monitoring, with three studies examining only baseline testing (Simon *et al.*, 2005; Lafata *et al.*, 2007; Besançon *et al.*, 2008). Six studies (Rhodes, 1992; Kalra *et al.*, 1999; Clayton *et al.*, 2006b; Sauer *et al.*, 2006; McAlister *et al.*, 2007; Raebel *et al.*, 2010) reported data on the both the proportion of patients with baseline biochemical testing and follow-up monitoring therapy, of which one (Clayton *et al.*, 2006b) presented data on the proportion of patients with testing before and monitoring after the initiation of antihypertensive therapy.

2.4.3 Summary of the primary outcome measure

The proportion of patients with biochemical testing prior to the initiation of therapy varied by the time frame used, the antihypertensive drug class, and the serum concentration measured. From 17 to 81% of patients treated with antihypertensive drugs had a baseline biochemical test (Table 2.1). There was also a four-fold range in the proportion of patients with follow-up monitoring (20–79%) (Table 2.2). In five of the eleven studies examining follow-up monitoring, fewer than half the patients had any evidence of biochemical monitoring.

Table 2.1 – Proportion of patients with baseline testing (grouped by the definition of baseline testing)

Study	Antihypertensive drug treatment	What to test?	% of patients with baseline testing
3 months prior			
Kalra (1999)	ACE inhibitors	Creatinine	45%
6 months prior to and 14 days after			
Simon (2005)	ACE inhibitors	Creatinine or potassium	67%
Simon (2005)	AT-II receptor antagonists	Creatinine or potassium	72%
Simon (2005)	Diuretics	Creatinine or potassium	67%
Lafata (2007)†	ACE inhibitors, AT-II receptor antagonists	Creatinine and potassium	59%
Lafata (2007) †	Diuretics	Potassium	55%
6 months prior			
McAlister (2007)	Thiazide diuretics	Sodium or potassium	21%
McAlister (2007)	Thiazide diuretics	Creatinine	32%
McAlister (2007)	ACE inhibitor, AT-II receptor antagonist, beta-blocker, calcium-channel blocker	Sodium or potassium	23%
McAlister (2007)	ACE inhibitor, AT-II receptor antagonist, beta-blocker, calcium-channel blocker	Creatinine	36%

Study	Antihypertensive drug treatment	What to test?	% of patients with baseline testing
Besançon (2008)	Spironolactone and ACE inhibitors	Potassium and creatinine	30%
2 years prior			
Clayton (2006)	Thiazide diuretics	Sodium and/or potassium	17%
No definition			
Rhodes (1992)	Diuretics	Urea or electrolytes	11%
Kalra (1999)	ACE inhibitors	Creatinine	49%
Sauer (2006)	ACE inhibitors or AT-II receptor antagonist	Potassium and creatinine	81%

†Only data from the control arm in patients who are new medication users are presented

Table 2.2 – Proportion of patients with follow-up monitoring (grouped by the definition of follow-up monitoring)

Study	Antihypertensive drug treatment	What to test?	% of patients with follow-up monitoring
2 weeks after			
Sauer (2006)	ACE inhibitor or AT-II receptor antagonist	Potassium and creatinine	27%
4 weeks after			
Sauer (2006)	ACE inhibitor or AT-II receptor antagonist	Potassium and creatinine	33%
3 months after			
Kalra (1999)	ACE inhibitors	Creatinine	29%
Sauer (2006)	ACE inhibitors or AT-II receptor antagonist	Potassium and creatinine	50%
1 year after			
Hurley (2005)	ACE inhibitors	Creatinine	65%
Hurley (2005)	ACE inhibitors	Potassium	60%
Raebel (2007b)	ACE inhibitors	Creatinine and potassium	68%
Raebel (2007b)	AT-II receptor antagonists	Creatinine and potassium	74%
Géradin-Marais (2008)	Diuretics	Serum chemistry monitoring	75%
Hurley (2005)	Diuretics	Creatinine	64%

Study	Antihypertensive drug treatment	What to test?	% of patients with follow-up monitoring
Hoch (2003)	Diuretics	Potassium	79%
Hurley (2005)	Diuretics	Potassium	66%
Rhodes (1992)	Diuretics	Urea or electrolytes	20%
Raebel (2007a)	Spironolactone	Creatinine and potassium	72%
McAlister (2007)	Thiazide diuretics	Sodium or potassium	38%
McAlister (2007)	Thiazide diuretics	Creatinine	41%
McAlister (2007)	ACE inhibitor, AT-II receptor antagonist, beta-blocker, calcium-channel blocker	Sodium or potassium	31%
McAlister (2007)	ACE inhibitor, AT-II receptor antagonist, beta-blocker, calcium-channel blocker	Creatinine	42%
2 years after			
Clayton (2006)	Thiazide diuretics	Sodium and/or potassium	32%
No definition			
Kalra (1999)	ACE inhibitors	Creatinine	62%
Raebel (2010)	ACE-inhibitors, angiotensin-II receptor blockers, spironolactone	Potassium	71%

2.4.4 Summary of secondary outcomes measures

Eight studies presented information on patient characteristics associated with either baseline testing or follow-up monitoring. Biochemical testing prior to the initiation of treatment was less likely to occur in women and in patients with fewer co-morbidities (Simon *et al.*, 2005). Few studies presented additional information besides the proportion of patients with baseline testing or follow-up monitoring.

Increasing age (Clayton *et al.*, 2006b; Raebel *et al.*, 2007a; Raebel *et al.*, 2007b; Gérardin-Marais *et al.*, 2008; Raebel *et al.*, 2010), more concomitant prescriptions

(Clayton *et al.*, 2006b; Raebel *et al.*, 2010), and increasing number of co-morbidities such as diabetes mellitus (Raebel *et al.*, 2007a; Raebel *et al.*, 2007b) or chronic renal failure (Raebel *et al.*, 2007b; Raebel *et al.*, 2010) have been shown to be significantly associated with follow-up monitoring. One study demonstrated that patients prescribed thiazide diuretics were more likely to have their serum electrolytes monitored than those prescribed angiotensin-converting enzyme (ACE) inhibitors, angiotensin-II (AT-II) receptor antagonists, or calcium-channel blockers (McAlister *et al.*, 2007). Patients with fewer outpatient encounters and those travelling longer distances to receive treatment were significantly less likely to have follow-up monitoring (Sauer *et al.*, 2006). Some studies have demonstrated that male patients were significantly more likely to have follow-up monitoring (Clayton *et al.*, 2006b; Raebel *et al.*, 2007a; Raebel *et al.*, 2007b). However, two studies found that female patients were likely to be monitored (Gérardin-Marais *et al.*, 2008; Raebel *et al.*, 2010).

Two studies presented additional information on the number of monitoring tests, demonstrating a low frequency of monitoring subsequent to the initial first test. Clayton and colleagues (2006b) provided data on the number of follow-up electrolyte tests, in addition to data on the proportion of patients with follow-up monitoring. McAlister and colleagues (2007) presented the test density (number of tests per 100 patients per 6 months) by drug class, with elderly patients treated with thiazide diuretics having the greatest density of tests.

One study examined the relationship between monitoring and adverse patient outcomes. In a sub-group of diabetic patients treated with antihypertensive therapy,

serum potassium monitoring was associated with a decreased risk of hyperkalaemia and hyperkalaemia-associated adverse events such as emergency department visits, hospitalization and death (Raebel *et al.*, 2010).

2.5 Critical appraisal of prior studies of monitoring of patients treated with antihypertensive drugs

This systematic review identified a range of studies examining the nature of monitoring for ADRs in patients treated with antihypertensive drugs. The estimated rate of baseline biochemical testing in primary care varied markedly, ranging from 17% to 81% (Rhodes, 1992; Kalra *et al.*, 1999; Simon *et al.*, 2005; Clayton *et al.*, 2006b; Sauer *et al.*, 2006; Lafata *et al.*, 2007; McAlister *et al.*, 2007; Besançon *et al.*, 2008). Similarly, the proportion of patients with follow-up monitoring ranged from 20% to 79% (Rhodes, 1992; Kalra *et al.*, 1999; Hurley *et al.*, 2005; Clayton *et al.*, 2006b; Sauer *et al.*, 2006; McAlister *et al.*, 2007; Raebel *et al.*, 2007a; Raebel *et al.*, 2007b; Raebel *et al.*, 2010). The wide range in the proportion of patients with baseline testing or follow-up monitoring may reflect differences in monitoring rates between studies, but may also be a result of differences in the methods and definitions of monitoring used by the various studies. These differences, and the various study populations, make comparison between studies challenging. Each study was assessed against several quality and methodological indicators, which are presented in Table 2.3 and summarized in Figure 2.2.

Table 2.3 – Summary of quality and methodological indicators

Study	Representativeness of the cohort?	Ascertainment of exposure to treatment?	Assessment of laboratory monitoring?	Validation of laboratory monitoring?	Length of follow-up sufficient?	Reference made to published monitoring guideline?	Monitoring assessed using insurance database?	Cohort of newly treated patients?	Baseline testing and follow-up monitoring presented?	Frequency of monitoring presented?	Patient factors associated with monitoring presented?	Relationship between monitoring and adverse outcomes presented?
Rhodes (1992)	+	+	–	–	+	–	–	–	+	–	–	–
Kalra (1999)	+	+	?	–	+	+	–	–	+	–	–	–
Hoch (2003)	?	+	+	–	+	–	–	–	–	–	–	–
Hurley (2005)	+	+	+	–	+	–	+	–	–	–	–	–
Simon (2005)	–	+	+	–	+	+	+	–	–	–	+	–
Clayton (2006b)	+	+	+	–	+	–	–	–	+	–	+	–
Sauer (2006)	?	+	?	–	+	+	+	–	+	–	+	–
Lafata (2007)	+	+	+	–	+	+	+	+	–	–	–	–
McAlister (2007)	–	+	+	–	+	+	+	+	+	+	+	–

Study	Representativeness of the cohort?	Ascertainment of exposure to treatment?	Assessment of laboratory monitoring?	Validation of laboratory monitoring?	Length of follow-up sufficient?	Reference made to published monitoring guideline?	Monitoring assessed using insurance database?	Cohort of newly treated patients?	Baseline testing and follow-up monitoring presented?	Frequency of monitoring presented?	Patient factors associated with monitoring presented?	Relationship between monitoring and adverse outcomes presented?
Raebel (2007a)	+	+	+	–	+	+	+	–	–	–	+	–
Raebel (2007b)	+	+	+	+	+	+	+	–	–	–	+	–
Besançon (2008)	+	+	+	–	+	+	+	–	–	–	–	–
Géradin-Marais (2008)	–	+	+	–	+	+	+	–	–	–	+	–
Raebel (2010)	–	+	+	–	?	+	+	+	+	–	+	+

+ Yes – No ? Unclear (insufficient information available to make a judgement)

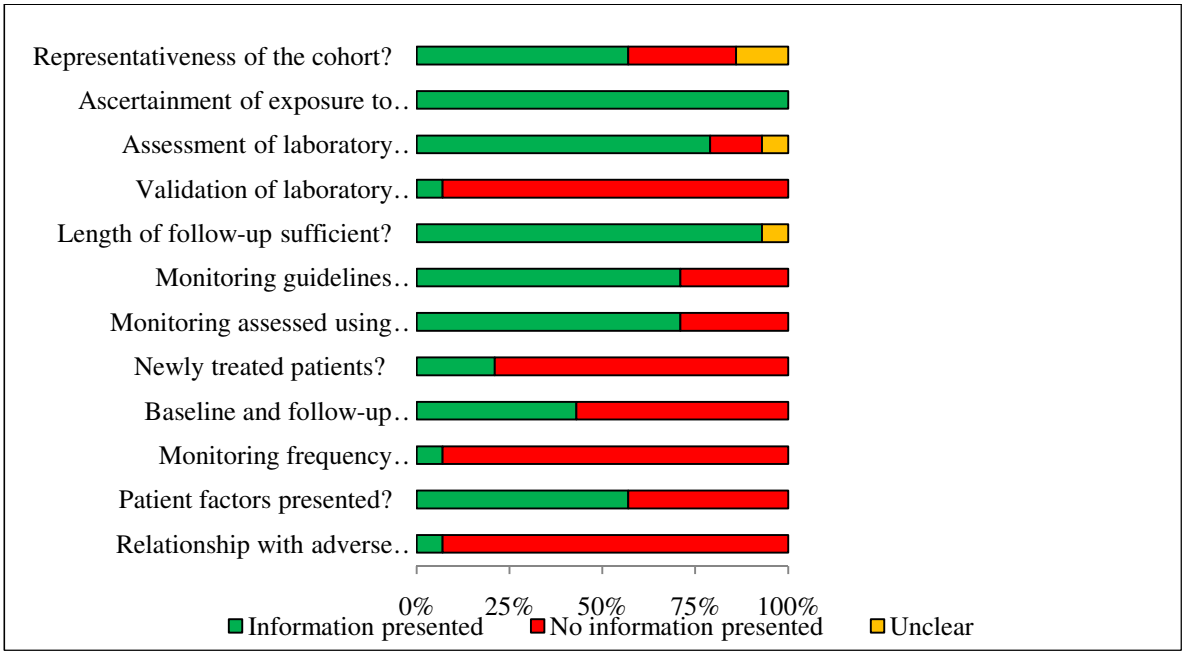


Figure 2.2 – Percentage of included studies where information for the methodological indicators are presented

Patients identified by the review differed in the extent to which they were representative of the patients who would receive biochemical monitoring due to treatment with antihypertensive therapy. While the majority of studies included patients as young as 18 years old, three were carried out exclusively in older patients (Simon *et al.*, 2005; McAlister *et al.*, 2007; Gérardin-Marais *et al.*, 2008). These three studies demonstrated high rates of monitoring, which may be due to doctors targetting monitoring to patients at higher risk of ADRs as adverse reactions to drug treatment have been shown to be more frequent in older patients. Three studies (Lafata *et al.*, 2007; McAlister *et al.*, 2007; Raebel *et al.*, 2010) presented data on monitoring in newly treated patients, who may present with different characteristics (e.g. co-morbidities, severity of hypertension, demographics) than patients who have been treated for some time. This has important implications for monitoring as patients who are newly treated may be particularly vulnerable to ADRs as they are drug-naïve and some ADRs, such as hyponatraemia, have been shown to occur very soon after the

initiation of treatment (Spital, 1999; Aronson, 2006). The sample size of the different studies also differed considerably. Although one study was carried out in over 160000 patients (McAlister *et al.*, 2007), the majority of studies examined fewer than 10000 patients.

All of the studies used electronic prescription records to identify patients exposed to treatment with antihypertensive therapy. Electronic patient records were also used to identify both baseline testing and follow-up monitoring in the majority of the studies. However, one study only used a retrospective chart review (Rhodes, 1992) and two studies (Kalra *et al.*, 1999; Sauer *et al.*, 2006) did not provide clear descriptions on the methods used to identify biochemical monitoring. Only Raebel and colleagues (2007b) validated the ability of the electronic records to identify monitoring. They considered the actual patients records to be the gold standard, and compared the monitoring results obtained from the electronic records with the patient records for a random sample of patients, to determine the sensitivity and predictive values of the administrative data.

The majority of the studies identified by the review assessed the nature of monitoring using large insurance or administrative databases. The nature of laboratory monitoring in these large organizations may be significantly different from monitoring carried out in other countries and in other healthcare organizations.

All of the studies described biochemical monitoring as an important tool for identifying adverse reactions to treatment. However, over a quarter of the studies

made no reference to a published guideline on monitoring patients for adverse reactions to antihypertensive therapy.

Most of the studies focused exclusively either on baseline testing or follow-up monitoring. Therefore very few studies were able to examine the important relationship between baseline testing and monitoring following the initiation of treatment. The majority of studies only treated monitoring (either baseline testing or follow-up monitoring) as a binary outcome. Indeed, only two studies provided any additional information on the number of monitoring tests (Clayton *et al.*, 2006b; McAlister *et al.*, 2007), which limits any analysis on the nature of monitoring.

Several studies only presented the prevalence of monitoring and did not provide any information on the patient factors associated with monitoring, which limits the interpretation of their results (Hoch *et al.*, 2003; Hurley *et al.*, 2005). Finally, only one study (Raebel *et al.*, 2010) examined any relationships between the monitoring of patients and adverse outcomes. Monitoring is advocated as a way of preventing patient harm but little to no information was presented in previous studies on the relationship between monitoring and adverse patient outcomes.

Barriers to monitoring

Monitoring is recommended as a tool for the identification of potential adverse reactions to therapy. Evidence obtained from the studies identified by this systematic review suggests that as many as one in five patients do not obtain any follow-up monitoring during treatment with antihypertensive therapy. Doctors have described monitoring as a critical, albeit time-consuming, component of their practice (Goldman

et al., 2010). Several barriers to monitoring have been described including: lack of consensus in the guidelines, uncertain responsibility for monitoring, patient non-adherence, and absence of alerts or reminders to monitor.

Strategies for improving monitoring

Several initiatives have been recently developed to encourage biochemical monitoring, which sometimes apply to the healthcare practitioner and sometimes to the patient. Electronic laboratory monitoring alerts aimed at reminding doctors to undertake monitoring have been examined within a computerized medical records system (Hoch *et al.*, 2003; Feldstein *et al.*, 2006; Palen *et al.*, 2006; Matheny *et al.*, 2008; Lo *et al.*, 2009). Academic detailing, where doctors go through a face-to-face educational process on the importance of monitoring, has also been investigated as a tool to improve monitoring (Lafata *et al.*, 2007). Additional research has investigated the impact of an automated voice message to the patient (Feldstein *et al.*, 2006), and a pharmacy team outreach to the patient (Feldstein *et al.*, 2006) on monitoring.

Although some slight improvement in the rates of appropriate biochemical monitoring have been observed following the introduction of the interventions, most were not statistically significant. Certainly electronic alerts may not alter physician behaviour due to issues relating to alert fatigue. The results obtained from the studies investigating the various interventions is also limited as the majority of these studies have been undertaken in American healthcare systems and the results may not be generalizable to other environments.

2.6 Conclusions

This systematic review identified several large, well planned studies that examined the prevalence of monitoring and identified patient factors associated with greater rates of monitoring. However, very few studies examined both baseline testing and follow-up monitoring. The majority of studies demonstrated that monitoring is not being carried out in accordance with published guidelines. In only 45% of the studies examining follow-up monitoring did more than half of the patients have any monitoring. There also remains a significant gap in knowledge with respect to the patterns of monitoring of antihypertensive drugs for ADRs in newly diagnosed and newly treated hypertensive patients. Further research that accurately examines potential relationships between monitoring and adverse patient outcomes is important in order to determine the effectiveness and value of monitoring.

Chapter 3

DESCRIPTION OF THE STUDY COHORT

In this chapter I introduce the General Practice Research Database (GPRD) and describe its use in pharmacoepidemiological research. I explain the process by which the cohort of patients who were newly diagnosed with hypertension and newly treated with antihypertensive therapy was identified from the GPRD. Finally, I define the various patient covariates and outcomes that were used in the analysis of the database.

3.1 The use of large databases in pharmacoepidemiology

Pharmacoepidemiology has been defined as the study of the use of and the effects of drugs in large numbers of people (Strom, 2005). It is a relatively new applied field that provides a link between epidemiology and clinical pharmacology. Pharmacoepidemiological methods are often applied to large automated databases or routine data sources such as MediPlus (Germany, UK, France), GPRD (UK), Saskatchewan Health (Canada), Medicaid (US), and PHARMO (The Netherlands). These databases provide a way to link patient records with drug therapy and outcomes for large populations in both a time- and cost-effective manner. Because they have the

potential to provide very large sample sizes, automated databases are particularly useful when looking for outcomes such as uncommon (fewer than 1 in 100 patients) or rare (fewer than 1 in 1 000 patients) ADRs. The size and nature of these databases is advantageous compared with clinical trials, which are most often carried out in small, pre-selected groups of patients over short periods of time. Clinical studies are usually not specifically designed to identify ADRs, instead focusing on drug efficacy, and are often too small or too short in follow-up time to detect rare, or late or delayed ADRs. Finally, the therapy used in the treatment of patients and the way in which the therapy is administered and monitored differs significantly between clinical trials and 'real life' clinical practice. Therefore automated databases can provide both a large sample size to identify rarer or delayed outcomes and a more comprehensive understanding of drug treatment usage in clinical practice.

The use of such databases in pharmacoepidemiology has, however, been criticized. Information on important confounding patient characteristics is often not recorded and because exposure is usually defined as a prescription record, misclassification of exposure may occur. The majority of these large databases do not contain direct linkage to hospital databases containing information on discharge diagnoses and procedures performed, which would be extremely useful (Garcia Rodriguez & Perez Gutthann, 1998).

3.2 The General Practice Research Database

The General Practice Research Database (GPRD) is the world's largest computerized database of anonymized longitudinal patient records from general practice (General Practice Research Database, 2007). The GPRD was established in 1987 and contains

information on over 3.5 million patients, providing approximately 39 million person-years of data. The GPRD contains approximately 5.5% of the UK population and is broadly representative of the general UK population in terms of age, sex, and geographic distributions (Gelfand *et al.*, 2005). The database is currently operated and maintained by the Medicines and Healthcare products Regulatory Agency (MHRA). Over 460 practices currently contribute to the GPRD and contributing GPs are provided with guidelines that define what information should be recorded electronically for GPRD purposes, including: demographic and lifestyle information (e.g. smoking status); medical diagnoses; all prescriptions; events leading to withdrawal of a drug or treatment; referrals to hospitals; treatment outcomes; and laboratory test results (General Practice Research Database, 2007).

The GPRD uses Read terminology, which is a structured hierarchy of both medical and non-medical terms. The codes include categories for observations, symptoms, diagnoses, investigations, occupations, and administrative processes. In 1994 the Read system of coding replaced the Oxford Medical Information System (OXMIS) medical codes, which were loosely based on the eighth revision of the International Classification of Diseases. OXMIS codes still remain in historical records in the GPRD.

The GPRD collects information on medications prescribed in general practice. Data collected for each prescription include generic name, formulation, strength, GPRD drug dictionary code, and codes relating to the British National Formulary (BNF) classification system.

The GPRD provides flat text files—plain text files that usually contain one record per line—for patient demographic, consultation, test, immunization, referral, and therapy data, which are all linked through the use of a unique patient identification number. The practices are regularly checked for data quality and data duplication, and referential integrity (i.e. that references between data are valid and intact). Patients must have registration, event recording, age, and sex data to a certain standard. Practices that meet the GPRD data quality standards are considered up-to-standard (UTS). A UTS date is generated, which indicates when data coding by the practice complied with specific quality measures.

3.2.1 Validation of the General Practice Research Database

The quality and the completeness of the data recorded in the GPRD have a direct impact on the validity of any research undertaken using the database. Internal methods (e.g. validating a diagnosis by the presence of codes indicating specific symptoms/signs, prescriptions for disease-specific drugs or test results) and external methods (e.g. questionnaires sent to GPs; comparison with other national databases or statistics) have been used to assess the validity of diagnoses recorded in the GPRD. The majority of validation occurred through requests to GPs for further information. Several studies have demonstrated good agreement between the medical diagnoses recorded using the Read/OXMIS coding systems and the diagnosis in the patient's medical record (Herrett *et al.*, 2010; Khan *et al.*, 2010). Outcomes such as myocardial infarction (Jick *et al.*, 1996), chronic obstructive pulmonary disease (Soriano *et al.*, 2001), inflammatory bowel disease (Lewis *et al.*, 2002), and venous thromboembolism (Lawrenson *et al.*, 2000) have been well validated. The

completeness of data from referrals to specialists and information related to such visits has been also well established (Jick *et al.*, 1991; Jick *et al.*, 1992).

3.2.2 Limitations of the General Practice Research Database

The GPRD provides a rich source of well-validated data, but there are some limitations to the database. Research has demonstrated that data from many specialists, as well as events that occur in hospital, may not be fully captured in the patient's electronic record. Treatments given in hospital may also not be recorded. The information on important confounding factors such as smoking, alcohol use, weight, and height is also limited (Gelfand *et al.*, 2005). Other factors such as socio-economic status are not implicitly recorded, although a patient's socio-economic status can be inferred from the level of deprivation of the post code of the GP practice. Finally, GPs do not routinely collect information on each encounter, only consultations. Therefore there is less complete information for minor sequelae of chronic diseases (like an episode of breathing problems in asthma), although the presence of the disease itself is usually very well recorded (Lawrenson *et al.*, 1999).

3.3 Development of the study cohort

3.3.1 Ethics approval

The study protocol was approved by the Independent Scientific Advisory Committee (ISAC) for Medicines and Healthcare products Regulatory Agency (MHRA) database research [study reference number 06_096, issued 24 Jul 2007].

3.3.2 Database development

3.3.2.1 Inclusion criteria

Data were requested from the MHRA using the following inclusion criteria:

1. Patients with a recorded first prescription of one of the nine following drug classes in general practice between 1 January 2000 and 31 December 2003 (first prescription was defined as no prior prescription of any drug from the nine drug classes within the preceding year);
    - I. Angiotensin-converting enzyme (ACE) inhibitors
    - II. Angiotensin-II (AT-II) receptor antagonists
    - III. Calcium-channel blockers (Ca-channel blockers)
    - IV. Thiazide diuretics
    - V. Potassium-sparing or loop diuretics
    - VI. Alpha-adrenoceptor antagonists (alpha-blockers)
    - VII. Beta-adrenoceptor antagonists (beta-blockers)
    - VIII. Aldosterone antagonists/potassium-sparing diuretics
    - IX. Mixed class (e.g. beta-blocker and thiazide diuretic combination)
2. Patients with a diagnosis of hypertension, as identified by a Read/OXMIS code of hypertension (Appendix 7) or with three blood pressure measurements of greater than 160/100 (the threshold for initiation of antihypertensive treatment without other patient factors at the time) on or in the 365 days before the date of the first antihypertensive prescription;

3. Patients must have at least one year of available data prior to the first prescription of an antihypertensive drug;
4. Patients must be aged 18 years and older on the day of the first prescription.

The day of the first prescription of the antihypertensive drug is referred to as the *index date*.

Data were obtained from the MHRA as 38 flat text tiles linked by a unique patient identification number. Look-up files were also provided to allow for the linkage between data files (Table 3.1). Not all of the text files were deemed to be relevant to the aims of the study including those relating to asthma, diet, exercise, immunization, passive smoking, residence, and sleeping habits.

Table 3.1 – GPRD data and look-up files obtained from the MHRA

File name	Data description
Data files	
ADR	Allergy and intolerance information
Agencies	Information about health agency involvement
Alcohol	Details relating to alcohol use including the number of units of alcohol per week
Asthma	Data recorded via the asthma disease management structured data area
Blood pressure	Current and historic blood pressure records
BMI	Historic measurements of height, weight, and BMI
Clinical	All the medical history data entered on the GP practice system, including systems, signs, and diagnoses (split into 3 files due to size of data set)
Consultation	Data relating to the type of consultation as entered by the GP (split into 2 files due to size of data set)
Death administration	Death data
Diabetes	Data entered via the diabetes disease management structured data area
Diet	Data relating to the patient's diet
Exercise	Data relating to the patient's exercise pattern
Height	Patient height data
Historical registration	Current and historical registration details

File name	Data description
Immunization	Data on patient immunizations.
Maternity	Data entered via the maternity structure data area
Passive smoking	Information about whether the patient is exposed to passive smoking
Patient	Basic patient demographics and patient registration details
Practice	Practice registration details
Referral	Information involving patient referrals to external care centres such as hospitals
Residence	Information about the patient's residential arrangements
Sleeping patterns	Data about the patient's sleeping habits
Smoking	Current and historic smoking details
Status	Current and historic records for the patient health status
Test	Data on the type of test (e.g. biochemical investigation) and results (split into 3 files due to size of data set)
Therapy	Data relating to all prescriptions issued by the GP (split into 6 files due to size of data set)
Treatment compliance	Data for which level the patient complies with the treatment issued
Weight	Weight data
Look-up files	
Medical codes	Read/OXMIS codes and associated code name
Product codes	Product codes for treatments including drug name, dose, and British National Formulary chapter and header
Test codes	Test codes and associated test name
Dose conversion table	Data that convert the dose provided to a numeric dose (e.g. one tablet per day converted to 1)

3.3.2.2 Data cleaning

The GPRD is a large and well-validated database, but extreme and implausible values can still exist within the database. When determining the patient baseline covariates described below in section 3.3.3, the recorded values were assessed for plausibility. Impossible values (such as a weight of 1000 kg) were recoded as an error or excluded. In determining baseline values, if the implausible value was the value recorded closest to the index date, this record was excluded and the next closest record was used as the baseline value.

The test file was examined in order to identify biochemical laboratory tests that were not clinically possible. In consultation with clinicians, decisions were made to exclude implausible laboratory test results. For example, a record of a creatinine serum concentration of 1800 µmol/l was excluded from the analysis, because such a concentration would not have been clinically possible.

3.3.2.3 Exclusion of pregnant women

Hypertension may be diagnosed during pregnancy as a result of the pregnancy or following pre-existing hypertension. Women who were pregnant during the study period were excluded because of possible differences in the condition, the treatment of the patient, and how the GP monitors the patient for both drug efficacy and drug safety.

An algorithm was developed based on the work suggested by Hardy and colleagues (2004) for identifying pregnant women using Read/OXMIS codes. The algorithm was based on the Read/OXMIS codes that were representative of a pregnancy marker or a pregnancy outcome.

Types of pregnancy markers included:

1. Lab tests and procedures;
2. GP practice visits related to pregnancy;
3. Threatened abortion;
4. Abortion referral;
5. Obstetric hospitalization.

Types of pregnancy outcomes included:

1. Elective termination;
2. Fetal death;
3. Hydatiform mole/blighted ovum;
4. Live births;
5. Delivery outcome unclear;
6. Delivery booking;
7. Multi-fetus delivery.

The algorithm developed by Hardy and colleagues (2004) was designed to identify definite pregnancies and to accurately determine dates of conception. My goal was to identify women who *possibly* could have been pregnant during the study time period and therefore 804 women were excluded from my analysis because they had any one of the following conditions:

1. A pregnancy marker from 01/04/1999 to 31/12/2003;
2. A pregnancy outcome from 01/01/2000 to 30/09/2004;
3. An expected date of delivery in the maternity file from 01/01/2000 to 30/09/2004.

Patients who had a pregnancy marker from 01/01/2004 to 30/09/2004 and no pregnancy outcome were also examined. The decision was taken to exclude these 339 patients on the basis that they were potentially pregnant during the study time period.

Therefore 1143 women, who represented 3.0% of the female population, were excluded from subsequent analyses.

3.3.3 Baseline patient covariates

A variety of demographic, lifestyle, and clinical covariates for all patients were identified and are described below. GPs are encouraged to record 'lifestyle factors' such as smoking status, alcohol intake, and weight at least once every three years. Therefore a five-year tolerance was considered suitable in order to include the maximum amount of relevant data.

Some covariates such as diabetes mellitus were identified through the use of Read/OXMIS codes. The development of a list of appropriate Read/OXMIS codes for relevant patient characteristics was carried out under the guidance of Dr P S Gill, a GP with considerable experience in disease coding in general practice; and Dr J J Coleman and Prof R E Ferner, who both have significant clinical experience and expertise.

Table 3.2 – Baseline patient covariates extracted from the GPRD

Patient covariate	Description	Category (if used)
Demographics		
Age	Ages were recorded as the age in years on the day of the index date	<40 years 40–49 years 50–59 years 60–69 years 70–79 years 80–89 years 90–100 years
Sex	Male or female	
Socio-economic status (SES)	A socio-economic status score (scored out of 100) is defined by the GPRD using the Index of Multiple Deprivation, which is based on the GP practice postcode	SES score quintiles (0= least deprived; 4=most deprived)
Lifestyle factors		
Smoking	Patients had their smoking status classified using the 'smoking summary' table provided by the GPRD or through the use of relevant Read/OXMIS codes (Appendix 8). The smoking status closest to index date, within 5 years of the index date, was used.	Never smoker Ever smoker (current or ex-smoker)
Alcohol use	The number of units of alcohol per week recorded closest to the index date (within 5 years) was used as the measure of alcohol intake.	Alcohol use was categorized using quartiles, which generally fit into recommended guidelines by sex. Male: 0–3 units/week; 4–14; 15–21; >21 Female: 0–2 units/week; 3–7; 8–13; >14
Height	Single height measurements where available were recorded; for patients with 3 or more height measurements, the median height measurement was taken; for 2 measurements, the mean was taken.	

<table>
<tr><th>Patient covariate</th><th>Description</th><th>Category (if used)</th></tr>
<tr><td>Weight</td><td>Weight in kilograms closest to the index date (± 5 years) was used</td><td></td></tr>
<tr><td>Body mass index (BMI)</td><td>BMI was defined as the above weight (in kg) divided by the square of the height (in metres) (kg/m^2)</td><td>BMI was categorized using recognized WHO definitions
<table><tr><th>Category</th><th>Range (kg/m^2)</th></tr><tr><td>Underweight</td><td><18 kg/m^2</td></tr><tr><td>Normal</td><td>18–24.9</td></tr><tr><td>Overweight</td><td>25–29.9</td></tr><tr><td>Obese</td><td>≥30</td></tr></table></td></tr>
<tr><td>Co-morbidities
Diabetes</td><td>Any record of a Read/OXMIS code of diabetes (Appendix 9) on or the index date was indicative of the presence of diabetes mellitus.</td><td></td></tr>
<tr><td>Baseline blood pressure</td><td>Systolic blood pressure (SBP) and diastolic blood pressure (DBP) readings recorded closest to the index date (within 365 days) were used. Where 3 values were recorded on the same day the average of the last two readings was used.</td><td>Blood pressure categories were created using guidelines from the British Hypertension Society (Williams et al., 2004)*
<table><tr><th></th><th>SBP (mmHg)</th><th>DBP (mmHg)</th></tr><tr><td>Normal</td><td><140</td><td><90</td></tr><tr><td>Mild</td><td>140–159</td><td>90–99</td></tr><tr><td>Moderate</td><td>160–179</td><td>100–109</td></tr><tr><td>Severe</td><td>≥180</td><td>≥110</td></tr></table>*If the SBP and DBP fell into different categories, the higher value was taken for classification</td></tr>
</table>

3.3.4 Assessment of biochemical serum concentrations

It is important to differentiate between biochemical tests recorded prior to the start of treatment and those tests started after the patient has been treated with antihypertensive therapy. Tests prior to the initiation of therapy can discover

secondary causes of hypertension, while tests taken during treatment can identify changes in serum concentration resulting from adverse reactions to the therapy.

I defined baseline biochemical *testing* as a measurement of one or more of serum creatinine, urea, potassium or sodium concentrations between six months prior to the index date and two days following it, to allow for any delays in the uploading of laboratory test results to the GP practice computer (prior to 2003 the majority of laboratory data entry was not automated). Biochemical *monitoring* was defined as a measurement of one or more of these serum concentrations between three days and one year after the index date.

Abnormal serum concentrations were determined using standardized (95%) reference ranges provided with the tests. Standardized reference ranges were 56–122 μmol/l for creatinine, 2.6–7.2 mmol/l for urea, 135–146 mmol/l for sodium, and 3.5–5.1 mmol/l for potassium. Where such ranges were missing, they were inferred from the patient's sex and age. Concentrations outside the reference ranges for sodium or potassium concentrations, and above the reference ranges for creatinine or urea, were classified as abnormal.

3.3.5 Changes or alterations to antihypertensive treatment

I identified several changes to patients' antihypertensive treatment using the methods described below.

3.3.5.1 Dose changes

Changes in the dose of the initial antihypertensive therapy were identified using the dose conversion look-up file provided by the GPRD. This file allows for the conversion of the text codes to a numeric dose (e.g. 'take tablet BD' is converted to the value of 2). The identification of a dose change was carried out as more vigilant monitoring of biochemical serum concentrations and renal function after an increase in the dose of antihypertensive therapy has been suggested (Eccles *et al.*, 1998; Chobanian *et al.*, 2003; North East Essex Medicines Management Committee, 2007; Smellie *et al.*, 2007).

3.3.5.2 Co-prescription of drug therapies

Additional antihypertensive therapies are often added after first-line therapy has demonstrated limited effectiveness. However, the use of two or more drug therapies at the same time can create the risk of a potential drug interaction. For example, the use of an ACE inhibitor and spironolactone at the same time can increase the risk of hyperkalaemia and close laboratory monitoring of serum potassium concentrations has been recommended (Juurlink *et al.*, 2004). I defined co-prescription of two drug therapies as a record for the second drug prior to the expiry of the prescription for the first drug (Figure 3.1). The day the second therapy was initiated was used as the day of co-prescription. The three drug groups identified, and the potential adverse reactions associated with their co-prescription, are described in Table 3.3.

Table 3.3 – Drug classes examined for co-prescription

Co-prescription of drug therapies	Potential ADR	Serum concentration test
ACE inhibitor & spironolactone	Hyperkalaemia	Potassium
ACE inhibitor & AT-II receptor antagonist	Hyperkalaemia	Potassium
Thiazide diuretic & loop diuretic	Hyperkalaemia	Potassium

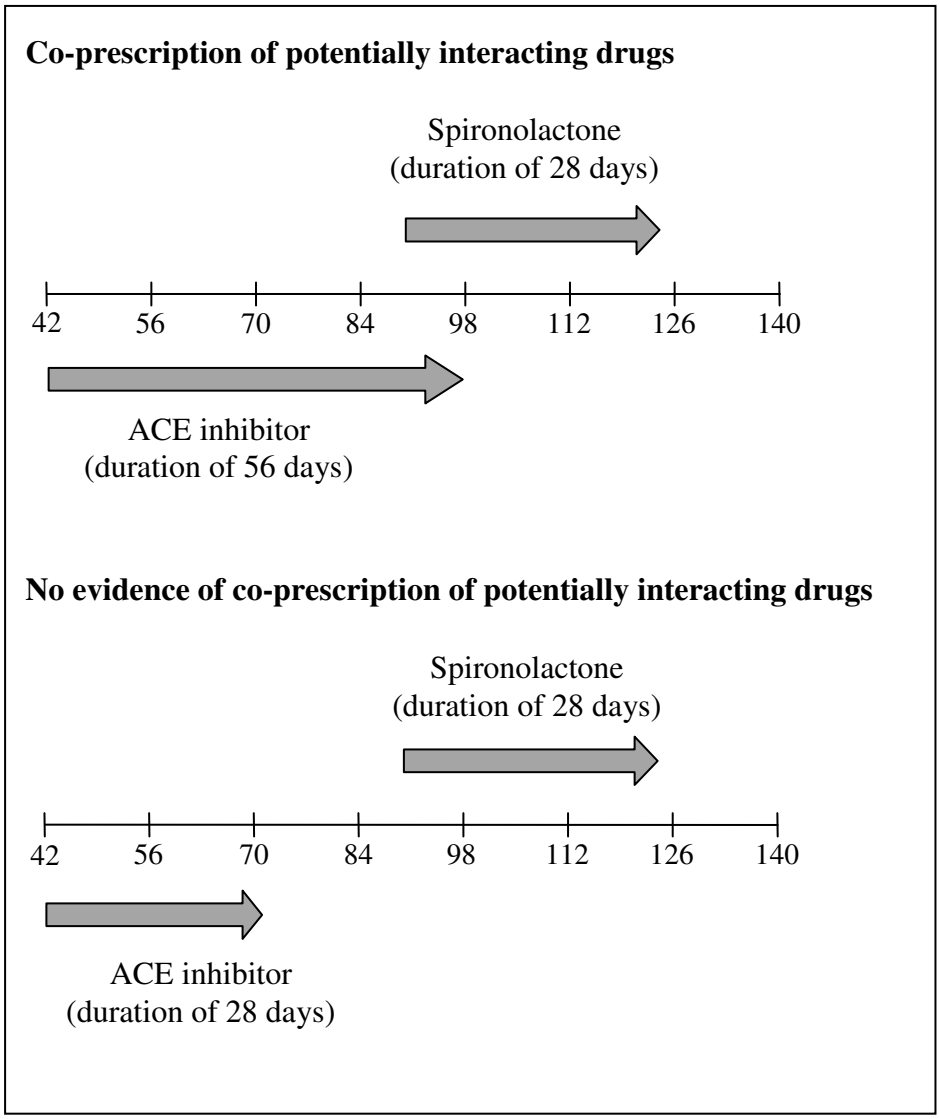


Figure 3.1 – Co-prescription was defined as the prescription of a second drug prior to the expiry of the prescription of the first drug

3.3.6 Patient outcomes

I identified four patient outcomes within one year of the first antihypertensive drug prescription, which are described below.

3.3.6.1 Persistence with antihypertensive drug treatment

Discontinuation of drug therapy is a common problem in clinical practice. Adherence to treatment is a term used to describe whether a patient has taken the prescribed treatment according to schedule, while persistence refers to whether a patient stays on therapy (Andrade *et al.*, 2006). Non-persistence can inhibit the treatment of certain chronic conditions and may lead to increased morbidity and mortality (van Wijk *et al.*, 2006). The lack of persistence may also suggest an adverse reaction that necessitated the cessation of treatment.

Several methods have been proposed for determining persistence with medication and there is no consensus about the best method for determining non-persistence in large automated databases (Caetano *et al.*, 2006). Some authors have used the medication possession ratio (MPR)—the proportion of days' supply obtained during a specific time period—to determine persistence. Persistence has also been defined using an 'anniversary' model; a patient is considered to be persistent with treatment for one year if they have a record of a prescription within a specified interval around the one-year anniversary of their first prescription. A minimum refill algorithm has also been used to defined persistence where records of a specified minimum number of prescriptions per year is indicative of persistence, although no consideration is given to the length or dates of the prescriptions. Finally, other papers have generally defined

discontinuation based on gaps or ‘grace periods’ between one dispensing of a drug and a subsequent prescription. Some studies have allowed for grace periods from 30–90 days after exhausting the drug supply from the prior prescription to define discontinuation (Bourgault *et al.*, 2005; Perreault *et al.*, 2005; Burke *et al.*, 2006; Elliott *et al.*, 2007).

I defined drug discontinuation as no subsequent prescription in the 30 days after the expiry of the prescription. The delay of 30 days was to allow for alternative supply and less than perfect dosing compliance (Figure 3.2). The discontinuation date was defined as the day when the previous prescription would have expired. Sensitivity analyses using alternative 60-day and 90-day periods (following the expiration of the most recent prescription) were carried out to determine the effect of different definitions of discontinuation.

I did not, however, take into account any drug stockpiling when determining the drug discontinuation date. The oversupply of therapy or ‘stockpiling’ can occur when a patient collects a prescription for later use (Greevy *et al.*, 2010), and this may have an impact on the calculation of drug discontinuation dates.

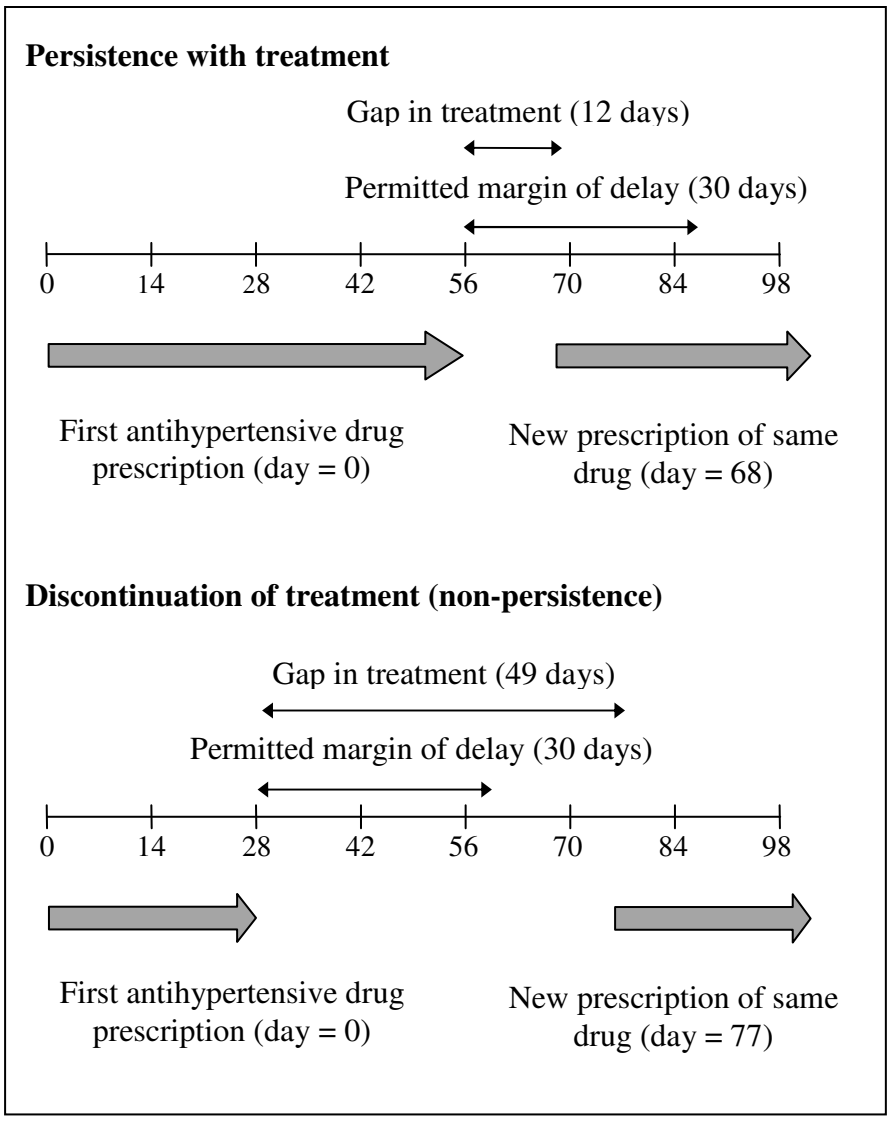


Figure 3.2 – Persistence with antihypertensive drug therapy. Discontinuation (non-persistence) was defined as no subsequent prescription in the 30 days after the expiry of the prescription

3.3.6.2 Biochemical adverse drug reactions

Biochemical ADRs (hyperkalaemia, hypokalaemia, and hyponatraemia) were identified using three methods: (1) a record of the ADR in the ADR database; (2) a Read/OXMIS ADR code (Appendix 10); or (3) a biochemical lab test above or below a certain concentration. Hyperkalaemia was defined as a serum potassium concentration greater than or equal to 5.5 mmol/l; hypokalaemia was defined as a

serum potassium concentration less than 3.5 mmol/l; and hyponatraemia was defined as a serum sodium concentration less than 135 mmol/l. The date closest to the index date was taken as the date of occurrence of the ADR.

The majority of cases of biochemical ADRs were identified solely on the basis of abnormal serum electrolyte concentrations: 87.8% of the hyponatraemia cases; 93.8% of the hypokalaemia cases; and 96.1% of the hyperkalaemia cases. A small proportion of the biochemical ADRs had both a record of an abnormal serum concentration and a Read/OXMIS code (2.5–6.8%) recorded within 28 days of each other.

3.3.6.3 Death

Death was identified through a combination of three methods suggested by researchers at the GPRD: (1) a Read/OXMIS code of death or suicide (Appendix 11); (2) a record of transfer out of the practice where the reason for transfer was death; or (3) a record in the GPRD 'Death administration' table.

During the data cleaning process, 19 patients who had been identified as dead within one year of antihypertensive treatment using the method described above, were found to have records after the recorded death date. A check was made of the clinical and therapy records for each of the 19 patients. All of these patients had clinical or therapy records after the reported date of death, which would suggest that the patients were indeed alive and the Read/OXMIS code had been entered in error. This may have been due to a death Read/OXMIS code being used for the description of the death of a spouse or family member and not death of the patient. Alternatively, some patients may have attempted suicide, which was recorded as a death in error. The decision was

made to recode these 19 patients as alive within one year of antihypertensive treatment.

3.3.6.4 Hospital admission

I wished to only identify admission to hospital and not simply hospital attendance. Therefore only Read/OXMIS codes where admission to hospital was clearly the outcome were included (Appendix 12). When two codes from the same patient were recorded within seven days of each other, the codes were treated as one single hospital admission, with the earlier date being used as the date of admission.

Chapter 4

BIOCHEMICAL LABORATORY MONITORING IN A COHORT OF PATIENTS TREATED WITH ANTIHYPERTENSIVE DRUGS: STATISTICAL METHODS

In this chapter I describe the statistical methods used to analyze the nature of biochemical laboratory monitoring in a cohort of patients treated with antihypertensive therapy. The results of these analyses are presented in Chapter 5. I characterize monitoring using the frequency, regularity, and responsiveness of biochemical laboratory tests. I introduce the concept of immortal time bias and how it can impact on the results of a study. Finally, I describe the statistical methods I have used to control for this type of bias.

4.1 Descriptive analyses

Descriptive analyses of the data were carried out to describe the patient cohort and to characterize biochemical laboratory monitoring. Comparisons were made between data using the Chi-squared test for categorical data, the independent t-test (for data

where parametric tests could be used), and the Mann-Whitney *U* test (for data where parametric tests were not appropriate) for continuous data.

I analyzed the data to determine the nature of biochemical monitoring by assessing the frequency, regularity, and responsiveness of laboratory testing, which are explained later. The analyses were based upon (1) the *number* of tests, (2) the calculated *time intervals* (Δt) between tests of serum concentration of creatinine, urea, sodium, and potassium, and (3) the *density* of tests (defined as the number of tests within 28 days). The density of tests was calculated in order to take into account the time that the patient had been in the study. For example, a patient who only had one test but had only been registered in the GP practice for two months would have approximately the same density of tests as someone who had six tests and had been registered in the practice for an entire year.

I assessed the frequency of biochemical monitoring by examining the number of tests, the density of tests, and the time intervals between the tests. Sparse monitoring occurs when there is a low number of biochemical tests, a low density of tests, or wide time intervals between individual tests. Conversely, frequent monitoring occurs when there is a high number of tests, a high density of tests, or a narrow separation in the time between tests (i.e. Δt is small).

I investigated the *regularity* of monitoring by examining the absolute first and second differences in the time intervals between biochemical tests (Figure 4.1).

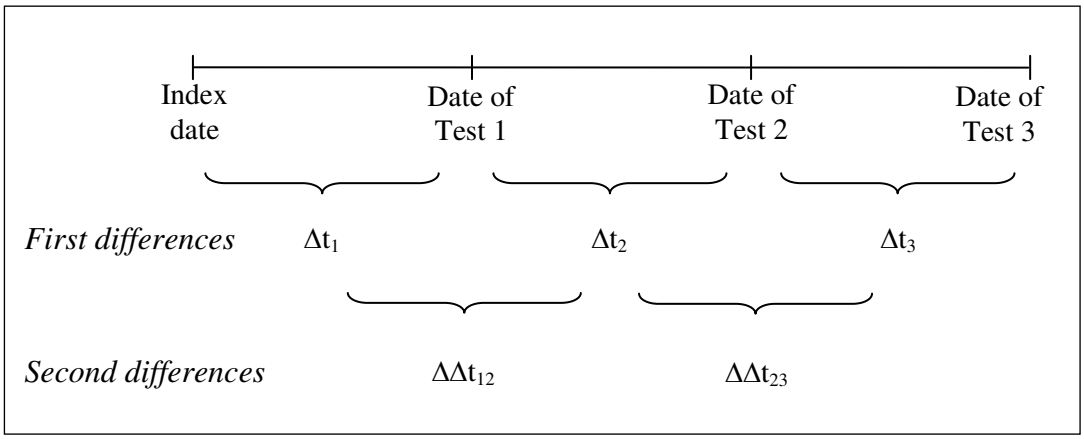


Figure 4.1 – First and second differences in the number of days between biochemical monitoring

Regular monitoring is characterized by approximately equal time intervals between biochemical tests (i.e. $\Delta t_1 \approx \Delta t_2 \approx \Delta t_3 \approx \Delta t_4$), while variable time intervals between tests is suggestive of irregular monitoring. The standard deviation for the time intervals between tests also provides an indication of the regularity of monitoring. A narrow standard deviation suggests a regular monitoring pattern, while a wide standard deviation in Δt suggests an irregular pattern of monitoring.

Regular monitoring can also be characterized through the examination of second differences. A second difference of 0 would indicate exactly regular monitoring. However, I chose to define regular monitoring as an average absolute second difference of 0 to 3 days, which would allow for any delays in uploading of the results or weekends. An average second difference of 0 to 7 days would suggest reasonably regular monitoring, while a value of greater than 7 days would suggest irregular

monitoring. Figure 4.2 demonstrates how the first and second differences were calculated and also provides an example of regular and irregular monitoring.

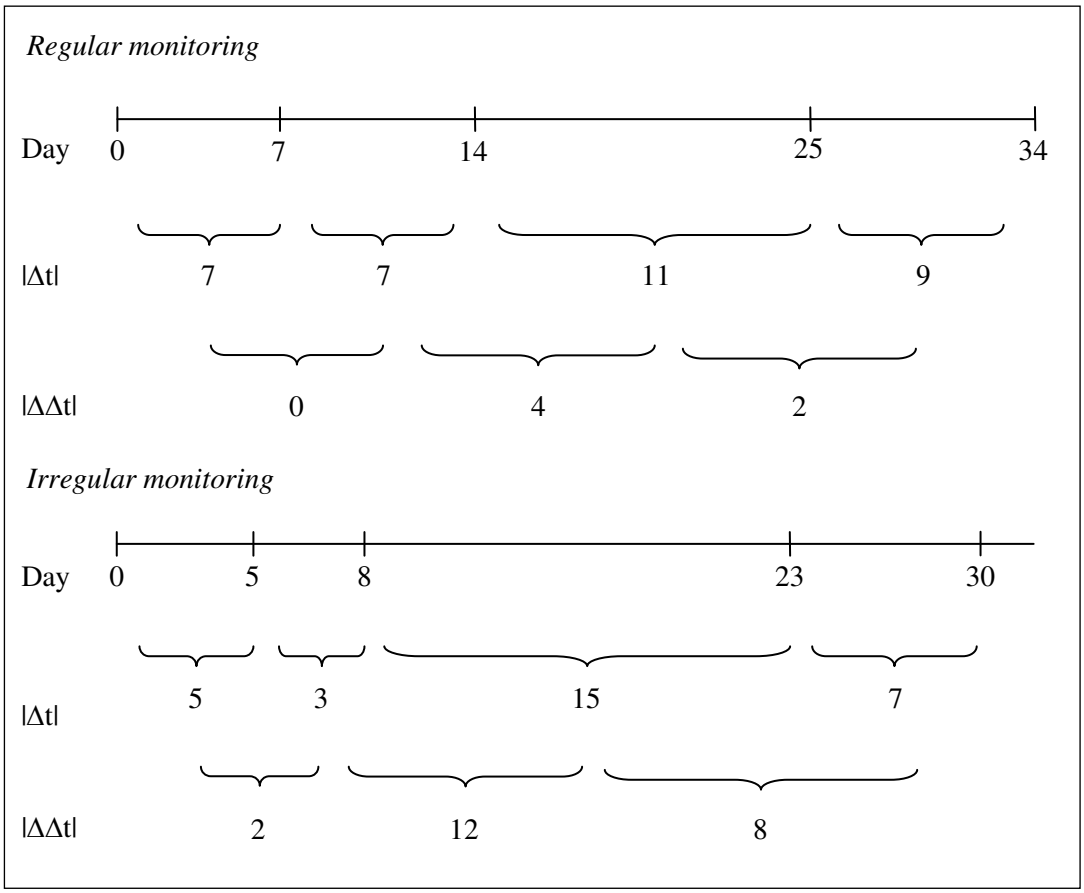


Figure 4.2 – Examples of regular and irregular monitoring based on first and second differences in the time intervals between biochemical monitoring

I assessed the *responsiveness* of monitoring to three events where guidelines have recommended an increase in the level of monitoring. First, I examined the nature of monitoring in patients where laboratory test results were outside the standardized 95% reference ranges provided with the test. Second, I looked at the relationship between monitoring and the co-prescription of other drugs that could lead to potentially harmful drug-drug interactions. Concomitant treatment with an ACE inhibitor and spironolactone can increase the risk of hyperkalaemia and close laboratory monitoring

of serum potassium concentrations has been recommended when the two drugs are used at the same time (Juurlink *et al.*, 2004). Finally, I assessed the responsiveness of monitoring to dose changes as some guidelines have specifically recommended more vigilant monitoring of serum electrolytes and renal function after an increase in the dosage of antihypertensive therapy.

4.2 Statistical analyses

4.2.1 Logistic regression modelling

I used Microsoft Access® to initially arrange and manipulate the data, and Stata® 10.0 to perform the statistical analyses. I determined the relationships between baseline patient characteristics and the probability of any monitoring within six months and one year in univariable logistic regression analyses. The methods by which the patient covariates were determined were described in Table 3.2. Alcohol use was excluded prior to undertaking any analyses due to the significant proportion of missing data. Baseline characteristics that were statistically significant at the $P<0.05$ level or were biologically plausible were then entered into a multivariable logistic regression model using backwards stepwise variable selection. The patient variable of sex was forced into the model *a priori*.

4.2.2 Time-to-event analyses

I used the Cox proportional hazards model to model the relationship between biochemical monitoring and patient outcomes (Cox, 1972). The Cox proportional hazards model is the most frequently used model for analyzing time-to-event data.

The Cox model is a regression method that provides an estimate of the ratio of the hazard function in treated patients compared with untreated patients. The hazard function is the probability that a patient experiences an event or outcome, conditional on the fact that they have not had an event or outcome to a certain time point. There are several advantages to using the Cox model including: (1) the ability to censor patients who fail to complete a study or do not reach the study endpoint; (2) the ability to incorporate time-varying patient covariates; and (3) the absence of any assumptions about the shape of the hazard over time (Cox & Oakes, 2001).

I modelled the relationship between biochemical monitoring and three patient outcomes described in section 3.3.6: death, hospital admission, and drug discontinuation. I also used the Cox proportional hazards model to model the relationship between evidence of hyponatraemia, hypokalaemia or hyperkalaemia and the same three outcomes.

Patients who were transferred out of the GP practice or who died were censored and the date of the censoring event was used as the follow-up time in the time-to-event analyses. A multivariable model using backwards stepwise variable selection was used, where patient characteristics that were statistically significant at the $P<0.05$ level or were biologically plausible were entered into the model. The patient variable of sex was also forced into the model *a priori*.

4.3 Immortal time bias

Biochemical follow-up monitoring was treated as a time-dependent covariate in the time-to-event analyses in order to avoid immortal time bias. This bias can occur when

an exposure or treatment of interest occurs within the same time period where an outcome can occur. Immortal time is the period of follow-up or observation time, during which an outcome (such as death) cannot occur (Suissa, 2007; Lévesque *et al.*, 2010). Immortal time bias has also been referred to as 'survivor treatment bias' (Austin & Platt, 2010), 'time-dependent bias'(Beyersmann *et al.*, 2008b), 'survival treatment selection bias' (Glesby & Hoover, 1996), or 'survival bias' (Zhou *et al.*, 2005).

This bias was first described in the early 1970s, when it was demonstrated in two cohort studies examining the benefit of heart transplantation (Messmer *et al.*, 1969; Clark *et al.*, 1971). Patients who received a heart transplant were shown to live longer than those who did not receive a heart transplant. A subsequent analysis demonstrated that the previous studies had included the time from when the transplant was approved to when the transplant was actually carried out in the total survival time for those patients who received a heart transplant (Gail, 1972). Therefore the time waiting for the transplantation had been incorrectly classified as time exposed to transplantation. This resulted in the time patients spent on the waiting list being incorrectly credited to the transplant and created an artificial decrease in the rate of death in the group of patients that did have a heart transplant.

4.3.1 The effect of immortal time bias

The effect of immortal time bias on the results of analyzes can be significant. A simulation study demonstrated that immortal time bias induced a negative or downward bias, which causes exposures or treatments to appear more protective than they really are (Austin *et al.*, 2006). Additional work by Beyersmann and colleagues

(2008a) showed that a log-hazard ratio that is biased due to immortal time bias will always be smaller than the true log-hazard ratio.

Several observational studies have demonstrated findings that have not been suggested or seen in randomized trials because of the failure to control for immortal time bias. For example, Rochon and colleagues (2000) demonstrated a 43% reduction in the rate of admission for heart failure following the use of low-dose beta-blockers while a 29% decrease in all-cause mortality was seen after treatment with inhaled corticosteroids (Sin & Tu, 2001).

Suissa (2007) clearly illustrated how immortal time bias can create an artificial protective effect when he examined the association between two drug classes with no plausible beneficial effect in the treatment of cardiovascular disease and the reduction in mortality. He found that when immortal time (the length of time from cohort entry to treatment) was excluded or misclassified, a 27% decrease in mortality was observed. When the immortal time was correctly classified, no significant difference in the association was found (Rate Ratio 0.94, 95% CI 0.73–1.20).
The effect of immortal time bias towards inducing a negative bias towards exposure or treatments has been further illustrated by studies that have re-analyzed published work that failed to account for immortal time bias. For example, Yee and colleagues (2004) presented the surprising result that treatment with 3-Hydroxy-3-methylglutaryl-coenzyme A (HMG Co-A) reductase inhibitors (‘statins’) reduced the risk of diabetes mellitus by 26%. Lévesque and colleagues (2010) re-analyzed the same data and correctly classified the time from cohort entry until prescription of a

statin as unexposed. They found an opposite effect, where exposure to a statin was in fact associated with a 53% increased risk of diabetes mellitus.

Another example focused on the study by Martin and colleagues (2006), which demonstrated a 51% reduction in the risk of all-cause mortality in patients with diabetes who performed self-monitoring of blood glucose. Hoffmann and Andersohn (2011) recently re-analyzed the same data set and correctly classified the immortal person-time as unexposed. They determined that self-monitoring of blood glucose was associated with a 95% increased risk of all-cause mortality, suggesting that the protective effect of monitoring was due entirely to immortal time bias.

A final example can be found in the study by Redelmeier and Singh (2001), which suggested that OSCAR® winners lived significantly longer (almost 4 years) than their less-recognized peers. A re-analysis of the study corrected for the winners' immortal time demonstrated that winner an Academy Award® conveyed an average advantage of only 0.7 years (95% CI –0.3 to 1.6 years) (Sylvestre *et al.*, 2006).

In summary, both simulation work and several re-analyses of published studies have demonstrated that the failure to control for immortal time can bias results and create the illusion of treatment effectiveness. When immortal time is misclassified, the event rate in the exposed group (the number of events per person-years) will be lower because the number of person-years in the exposed group is artificially increased. This will cause the rate of the event in the exposed group to be lowered. Therefore, a comparison between exposed and unexposed patients will be biased downwards and

exposed patients will be artificially protected until they are exposed to the treatment or intervention.

4.3.2 The extent of immortal time bias in the literature

Two studies have demonstrated that the failure to control for immortal time bias is prevalent in the medical literature. The review by van Walraven and colleagues (2004) examined 682 observational studies published in medical journals that used survival analysis. They identified 52 articles susceptible to immortal time bias and reported that in 44% of these studies, the use of analyses that corrected for this bias could have changed the study's conclusion.

A later paper identified twenty recently published studies, that all failed to take into account the risk of immortal time bias by either misclassifying or excluding the immortal time (Suissa, 2007). The majority of the studies used existing, large computerised databases and demonstrated significant decreases in the rates of adverse patient outcomes associated with treatment with various drug therapies.

4.3.3 Controlling for immortal time bias in the analysis of follow-up biochemical monitoring

A patient with follow-up biochemical monitoring was defined as any patient with a biochemical measurement in the 365 days following the start of antihypertensive treatment. The period from first antihypertensive prescription to the time of the first biochemical monitoring was considered to be event-free and thus 'immortal' (Figure 4.3).

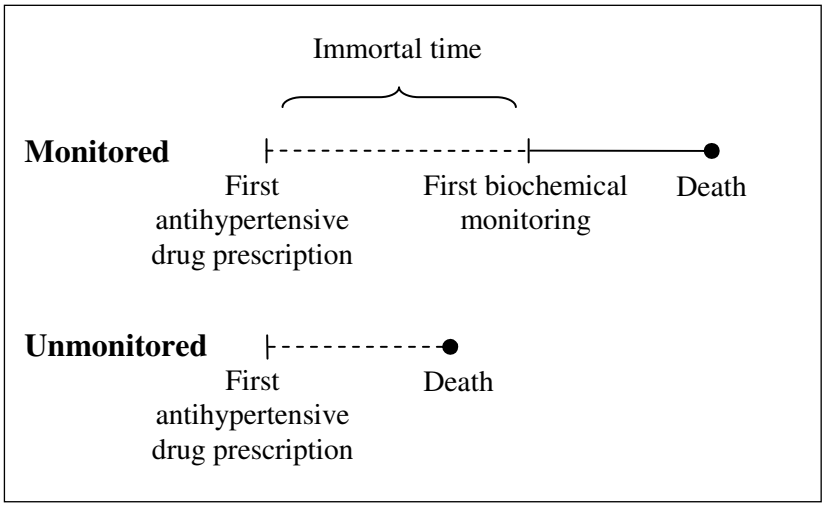
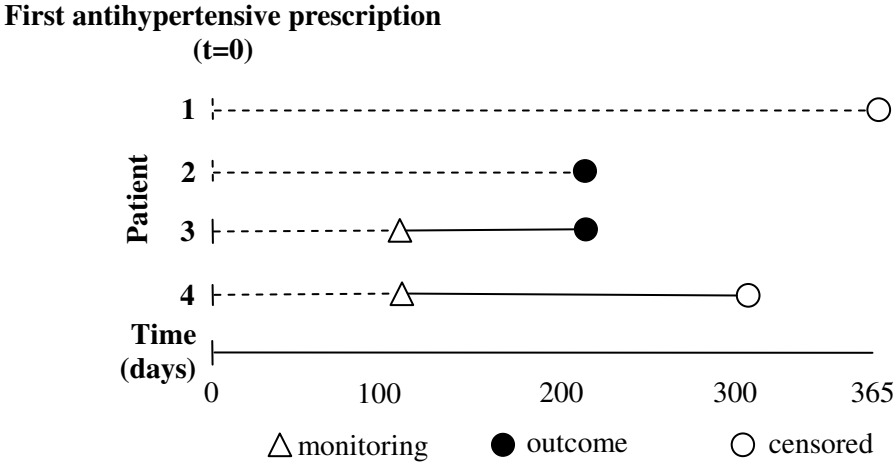


Figure 4.3 – Illustration of immortal time bias, using the adverse outcome of death as an example

For the patient to have been monitored, they would have needed to have survived this period of time. The failure to account for this time in the data analysis can lead to immortal time bias, which would bias the results in favour of those who were monitored.

Specifically, when an adverse patient outcome occurred shortly after the start of antihypertensive treatment, a patient was less likely to be classified as having biochemical monitoring because the opportunity for having a biochemical test was lower. Therefore a majority of outcomes that occurred shortly after antihypertensive treatment was initiated would have been classified as not monitored because there would have been fewer opportunities to receive biochemical monitoring. If the immortal time was not accounted for correctly, this would lead to a higher rate of the outcomes in the patients that did not have monitoring and would cause an artificial decrease in the rate of the outcome among patients that had biochemical monitoring.

Therefore, in order to control for immortal time bias I needed to model follow-up biochemical monitoring as a time-dependent covariate, which is illustrated in Figure 4.4. In this example, patient 3 had a record of follow-up monitoring on day 115 and experienced the outcome on day 211. The time from day 0 to day 115 is considered as 'immortal' and therefore was classified as unmonitored, while the time from day 115 to day 211 was classified as monitored.



Patient	Time	Time to event	Monitored?	Outcome?
1	0	365	0	0
2	0	211	0	1
3	0	115	0	-
3	115	211	1	1
4	0	115	0	-
4	115	310	1	0

Figure 4.4 – An illustration of how time was coded in order to control for immortal time bias

Finally, I wished to quantify the extent of misclassified immortal time bias and to estimate the effect of this bias on the relationship between biochemical monitoring and adverse patient outcomes. This was achieved through the use of a simple Poisson

rate approach, which assumed a constant rate of the various patient outcomes over follow-up time. I also used a Cox proportional hazards model to calculate hazard ratios adjusted for potentially confounding factors, but where the time to monitoring was time fixed (not time varying). Therefore the person days between entry into the cohort and biochemical monitoring were classified as monitored.

Chapter 5

BIOCHEMICAL LABORATORY MONITORING IN A COHORT OF PATIENTS TREATED WITH ANTIHYPERTENSIVE DRUGS: RESULTS

This chapter presents the results of the analysis investigating the nature and frequency of biochemical laboratory monitoring. I describe the frequency of both baseline testing and follow-up monitoring and make comparisons between patients with and without baseline biochemical testing. I demonstrate that follow-up monitoring, when carried out, is sparse and infrequent. I also show that few patients have both baseline biochemical testing and follow-up monitoring. Finally, I demonstrate that multiple patient characteristics are associated with follow-up monitoring and examine the relationships between monitoring and adverse patient outcomes.

5.1 Identification of the study cohort

I identified 77 905 patients from 401 GP practices who were newly diagnosed with hypertension and newly treated with an antihypertensive drug between January 2000 and December 2003. The majority of patients (69%) were identified based solely on a

recorded OXMIS/Read code for hypertension, while almost one third (28%) were selected for inclusion because they had both a code for hypertension and blood pressure measurements indicating hypertension. Only 3% of patients were identified solely on blood pressure measurements. 1143 women were excluded from the analysis because they could have been pregnant during the treatment period (Figure 5.1).

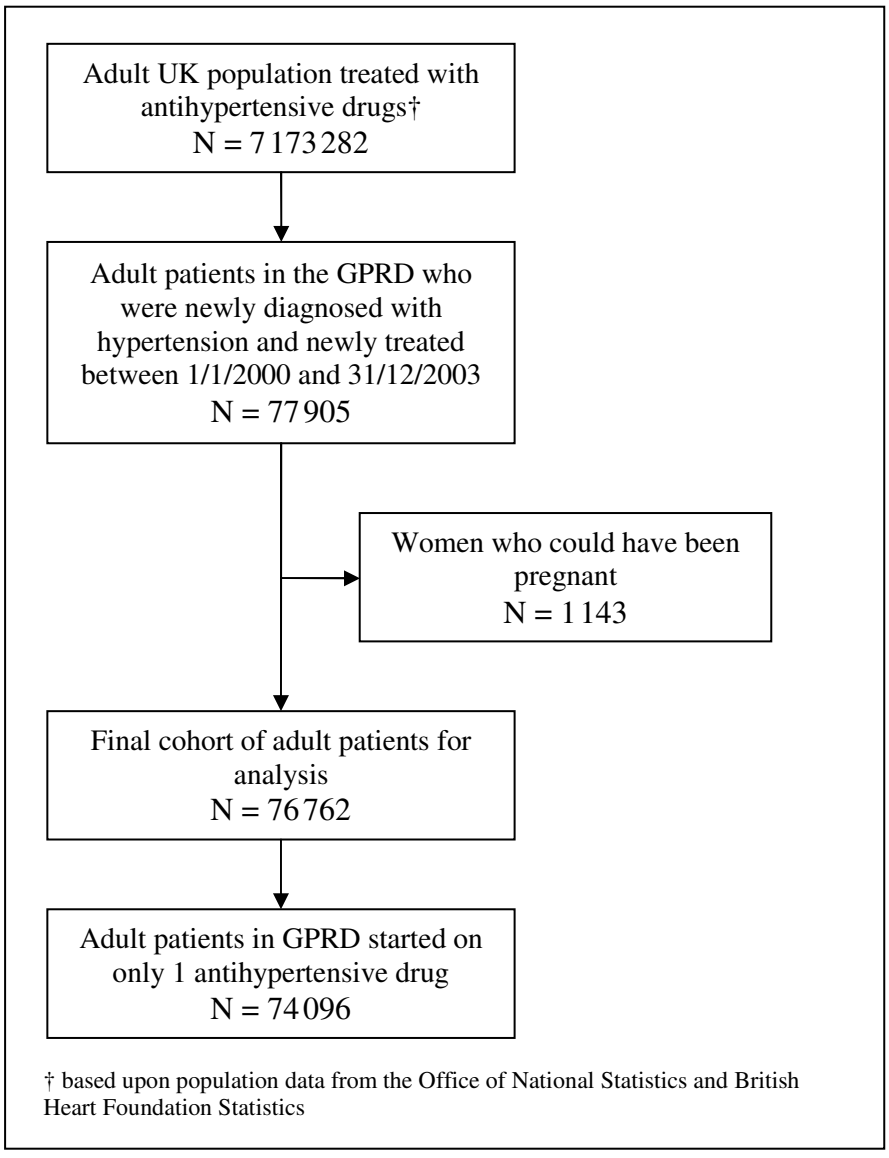


Figure 5.1 – Flowchart of patients treated with antihypertensive drugs

74 096 patients, 48% male, were first prescribed a single antihypertensive agent (the index prescription) between January 2000 and December 2003. Patients most often began treatment with a thiazide diuretic (48.1%), beta-blocker (24.3%), or an ACE inhibitor (15.2%). Although the majority of patients were treated with a thiazide diuretic, this proportion decreased from 2000 to 2002 and later increased in 2003 (Table 5.1). The opposite trend was found with ACE-inhibitor treatment where there was an increase in the proportion of patients treated from 2000 to 2002, followed by a significant decrease.

Table 5.1 – Choice of antihypertensive drug class by year of first prescription

Drug class	**Total (N=74 096)**	**Year of first antihypertensive prescription**				
		2000 (N=16 508)	**2001 (N=17 927)**	**2002 (N=19 532)**	**2003 (N=20 129)**	**P value†**
ACE inhibitor	**11 245 (15.2)**	2 168 (13.1)	2 682 (15.0)	3 319 (17.0)	3 076 (15.3)	<0.0005
Alpha-blocker	**753 (1.0)**	200 (1.2)	233 (1.3)	177 (0.9)	143 (0.7)	<0.0005
AT-II receptor antagonist	**1 630 (2.2)**	254 (1.5)	349 (2.0)	512 (2.6)	515 (2.6)	<0.0005
Beta-blocker	**17 977 (24.3)**	4 332 (26.2)	4 657 (26.0)	4 856 (24.9)	4 132 (20.5)	<0.0005
Ca-channel blocker	**5 651 (7.6)**	1 454 (8.8)	1 407 (7.9)	1 417 (7.3)	1 373 (6.8)	<0.0005
Combination	**771 (1.0)**	196 (1.2)	196 (1.1)	198 (1.0)	181 (0.9)	0.046
Loop or K-sparing diuretic	**416 (0.6)**	121 (0.7)	107 (0.6)	97 (0.5)	91 (0.5)	0.002
Thiazide diuretic	**35 653 (48.1)**	7 783 (47.2)	8 296 (46.3)	8 956 (45.9)	10 618 (52.8)	<0.0005

†Test for linear trend by year

Table 5.2 – Choice of antihypertensive drug class by age group at index date

Drug class	Age group						
	<40 (N=3 492)	**40–49 (N=10 388)**	**50–59 (N=20 360)**	**60–69 (N=19 312)**	**70–79 (N=14 971)**	**80–100 (N=5 573)**	**P value†**
ACE inhibitor	674 (19.3)	1 933 (18.6)	3 400 (16.7)	2 907 (15.1)	1 846 (12.3)	485 (8.7)	<0.0005
Alpha-blocker	27 (0.8)	60 (0.6)	177 (0.9)	233 (1.2)	197 (1.3)	59 (1.1)	<0.0005
AT-II receptor antagonist	116 (3.3)	307 (2.9)	518 (2.5)	363 (1.9)	251 (1.7)	75 (1.3)	<0.0005
Beta-blocker	1 165 (33.4)	3 368 (32.4)	5 928 (29.1)	4 357 (22.6)	2 474 (16.5)	685 (12.3)	<0.0005
Ca-channel blocker	257 (7.4)	657 (6.3)	1 252 (6.2)	1 570 (8.1)	1 356 (9.1)	559 (10.0)	<0.0005
Combination	29 (0.8)	84 (0.8)	217 (1.1)	194 (1.0)	168 (1.1)	79 (1.4)	0.007
Loop or K-sparing diuretic	11 (0.3)	15 (0.1)	59 (0.3)	90 (0.5)	135 (0.9)	106 (1.9)	<0.0005
Thiazide diuretic	1 213 (34.7)	3 964 (38.2)	8 809 (43.3)	9 598 (49.7)	8 544 (57.1)	3 535 (63.4)	<0.0005

† Test for difference across age groups

There were also some substantial differences in the prescription of antihypertensive therapy by age group (Table 5.2). The prescription of ACE inhibitors declined with age, while the prescription of all of the diuretic classes increased with age. Some classes of drugs demonstrated a more complex pattern, where there was an initial increase with age, followed by a decline in the oldest age group.

5.2 Describing the nature of baseline biochemical testing

31 094 patients (42%) had evidence of any baseline biochemical testing, which was defined as one or more laboratory tests in the six months prior to the index date. The majority of patients (76.2%) with baseline testing had all four baseline tests, while 2090 (6.7%) patients had only one test. Almost 12% of the baseline tests were outside of the standardized 95% reference range and were therefore abnormal (Table 5.3; Figure 5.2). A large number of patients had a record of serum urea concentration of 2.5, which is most likely a reflection of a tendency to record concentrations at the lower limit of normal.

Table 5.3 – Summary of baseline biochemical tests

Baseline biochemical test	Number of patients with a baseline biochemical test	Number of patients with an abnormal baseline test
Creatinine, n (%)	30 258 (40.8)	1 222/30 258 (4.0)
Urea, n (%)	25 263 (34.1)	93/25 263 (0.3)
Sodium, n (%)	28 262 (38.1)	969/28 262 (3.4)
Potassium, n (%)	28 061 (37.8)	1 627/28 061 (5.8)

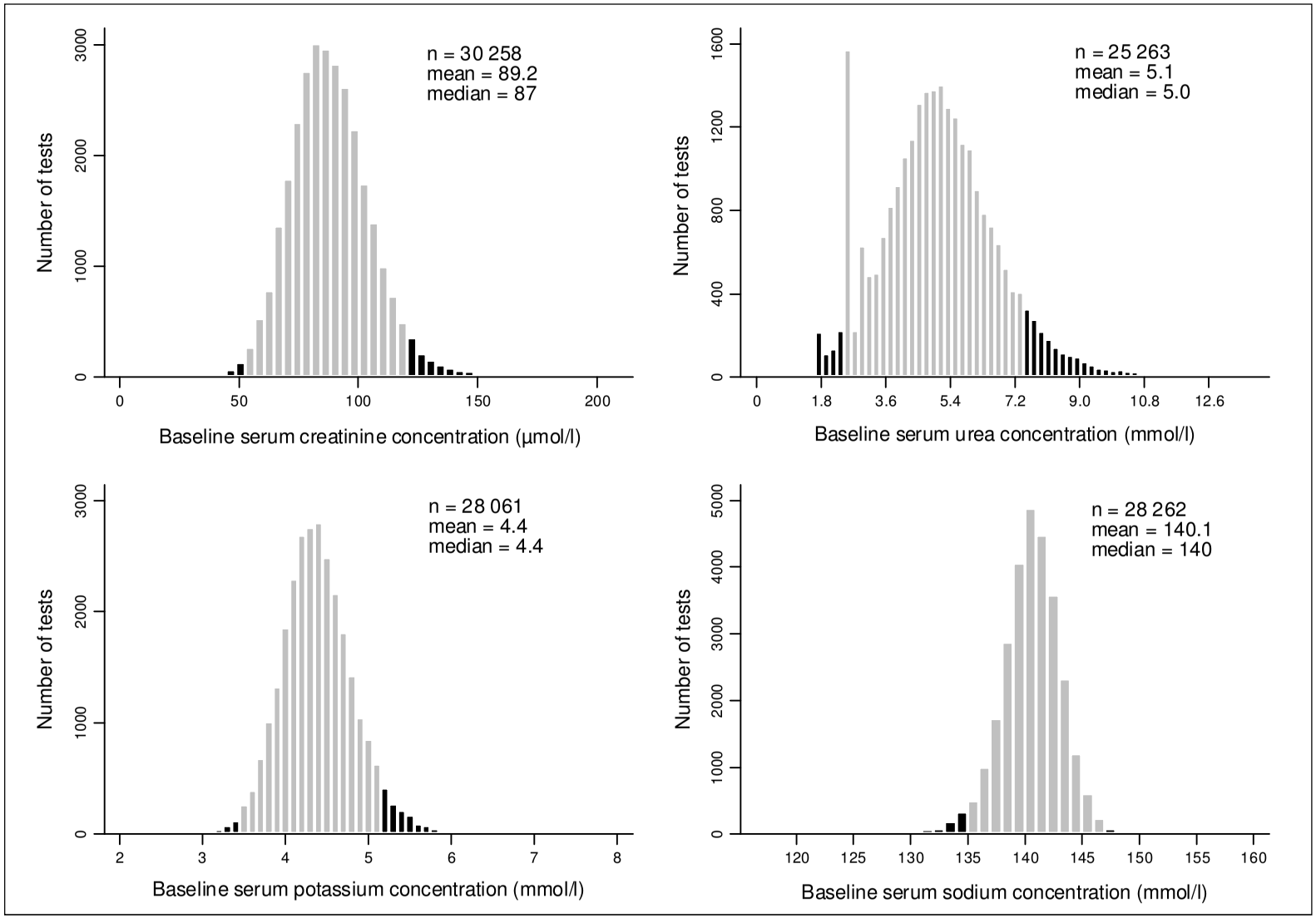


Figure 5.2 – Distribution of biochemical test results at baseline (black indicates concentrations outside standardized (95% CI) reference ranges)

There were significant differences between patients with and without baseline testing. Patients with baseline testing were more likely to be male, younger, have diabetes mellitus, and be a smoker (Table 5.4).

Table 5.4 – Baseline patient characteristics in those patients with baseline biochemical testing compared with those with no baseline biochemical testing

Baseline patient characteristic	Total (N = 74096)	Any baseline testing (N=31094)	No baseline testing (N=43002)	P value
Male, n (%)	35345 (47.7)	15513 (49.9)	19832 (46.1)	<0.0005
Age, mean (SD)	61.1 (12.8)	60.8(12.4)	61.2 (13.0)	0.0001
<40, n (%)	3492 (4.7)	1384 (4.5)	2108 (4.9)	<0.0005†
40–49, n (%)	10388 (14.0)	4396 (14.1)	5992 (13.9)	
50–59, n (%)	20360 (27.5)	8664 (27.9)	11696 (27.2)	
60–69, n (%)	19312 (26.1)	8429 (27.1)	10883 (25.3)	
70–79, n (%)	14971 (20.2)	6231 (20.2)	8740 (20.0)	
80–89, n (%)	5190 (7.0)	1875 (6.0)	3315 (7.7)	
90–100, n (%)	383 (0.52)	115 (0.37)	268 (0.62)	
BMI (kg/m^2), mean (SD)	28.5 (5.4)	28.6 (5.4)	28.4 (5.4)	0.0001
Underweight, n (%)	654 (1.0)	264 (1.0)	390 (1.1)	<0.0005†
Normal, n (%)	16039 (25.6)	6641 (24.6)	9398 (26.3)	
Overweight, n (%)	25392 (40.5)	11067 (41.0)	14325 (40.1)	
Obese, n (%)	20606 (32.9)	9004 (33.4)	11602 (32.5)	
Current or ex-smoker, n (%)	33118 (46.4)	14513 (48.1)	18605 (45.3)	<0.0005
Presence of diabetes mellitus, n (%)	6437 (8.7)	3876 (12.5)	2561 (6.0)	<0.0005
SES quintile, n (%)				
0 (least deprived)	15660 (21.1)	6132 (19.7)	9528 (22.2)	<0.0005†
1	13205 (17.8)	5875 (18.9)	7330 (17.1)	
2	15005 (20.3)	6121 (19.7)	8884 (20.7)	
3	14154 (19.1)	5311 (17.1)	8843 (20.6)	
4 (most deprived)	16072 (21.7)	7655 (24.6)	8417 (19.6)	
Blood pressure mmHg (SD)				
Mean systolic blood pressure	170.9 (20.1)	170.9 (19.5)	170.7 (20.5)	0.67
Mean diastolic blood pressure	97.2 (11.3)	97.2 (11.3)	97.2 (11.4)	0.62

Baseline patient characteristic	Total (N = 74096)	Any baseline testing (N=31094)	No baseline testing (N=43002)	P value
ACE inhibitor	11245 (15.2)	5835 (18.8)	5410 (12.6)	<0.0005
Alpha-blocker	753 (1.0)	299 (1.0)	454 (1.1)	0.21
AT-II receptor antagonist	1630 (2.2)	668 (2.2)	962 (2.2)	0.42
Beta-blocker	17977 (24.3)	7123 (22.9)	10854 (25.2)	<0.0005
Ca-channel blocker	5651 (7.6)	2229 (7.2)	3422 (8.0)	<0.0005
Loop or K- sparing diuretic	416 (0.6)	151 (0.49)	265 (0.62)	0.02
Combination preparation	771 (1.0)	246 (0.79)	525 (1.2)	<0.0005
Thiazide diuretic	35653 (48.1)	14543 (46.8)	21110 (49.1)	<0.0005

†Chi-squared test for trend

5.3 Follow-up monitoring

5.3.1 What is the nature of follow-up monitoring?

In the year after the date of first prescription half of the patients (37365) had at least one biochemical measurement (Table 5.5; Figure 5.3). Over a third of patients (26946; 36.4%) had at least one measurement within six months and 16.3% had at least one measurement within 28 days. Of those patients who were monitored at least once within one year, 29753 (79.6%) had all four tests of potassium, sodium, creatinine, and urea; 2045 (5.5%) of those who were monitored had only one type of serum concentration test.

Almost 40% of those patients who were monitored had at least a second biochemical serum concentration measurement. The majority of patients tended to have 'sparse' monitoring in that they had few, if any, biochemical tests and a low density of tests. 8790 (23.5%) patients with any monitoring had at least one serum concentration that

was abnormal and 6589 (74.9%) of these abnormal serum concentrations were recorded on the first follow-up monitoring test.

Table 5.5 – Summary of the biochemical serum concentrations measured in the year after the start of antihypertensive treatment

	Creatinine	Urea	Sodium	Potassium
Number of patients with at least 1 test				
Within 1 year, n (%)	36558 (49.3)	31318 (42.3)	34595 (46.7)	34606 (46.7)
Within 6 months, n (%)	26282 (35.5)	22456 (30.3)	24804 (33.5)	24853 (33.5)
Within 1 month, n (%)	11709 (15.8)	10124 (13.7)	11134 (15.0)	11222 (15.2)
Number of patients with				
1 test, n (%)	21899 (29.6)	19030 (25.6)	20819 (28.1)	20947 (28.3)
2 tests, n (%)	8840 (11.9)	7466 (10.1)	8338 (11.3)	8288 (11.2)
3 tests, n (%)	3454 (4.7)	2870 (3.9)	3260 (4.4)	3228 (4.4)
4 tests, n (%)	1370 (1.8)	1132 (1.5)	1258 (1.7)	1253 (1.7)
≥5 tests, n (%)	995 (1.3)	820 (1.1)	920 (1.2)	890 (1.2)
Number of tests per patient (in patients with at least 1 test) (mean; mode; range)				
Within 1 year	1.68 (1) 1–30	1.67 (1) 1–18	1.68 (1) 1–30	1.67 (1) 1–30
Within 6 months	1.39 (1) 1–15	1.38 (1) 1–15	1.39 (1) 1–15	1.38 (1) 1–15
Within 1 month	1.04 (1) 1–4	1.05 (1) 1–4	1.05 (1) 1–4	1.05 (1) 1–4
Density of tests per month, mean (SD)	0.064 (0.090)	0.054 (0.084)	0.060 (0.088)	0.060 (0.089)
Density of tests per month (in pts with at least one test), mean (SD)	0.129 (0.090)	0.128 (0.087)	0.129 (0.089)	0.128 (0.089)
Number of patients with abnormal test, n (%)	2654/36558 (7.3)	5098/31318 (16.3)	2894/34595 (8.4)	4547/34606 (13.1)
Number of patients with abnormal first test, n (%)	2220/36558 (6.1)	4397/31318 (14.0)	2115/34595 (6.1)	2898/34606 (8.4)

In the 1995 patients who had a potassium test subsequent to an abnormal test, 71% (1420) of the patients' serum potassium concentrations returned to normal. This was more than the 55% (750/1368), 36% (640/1794), and 36% (1312/3618) of patients whose sodium, creatinine, and urea tests returned to normal.

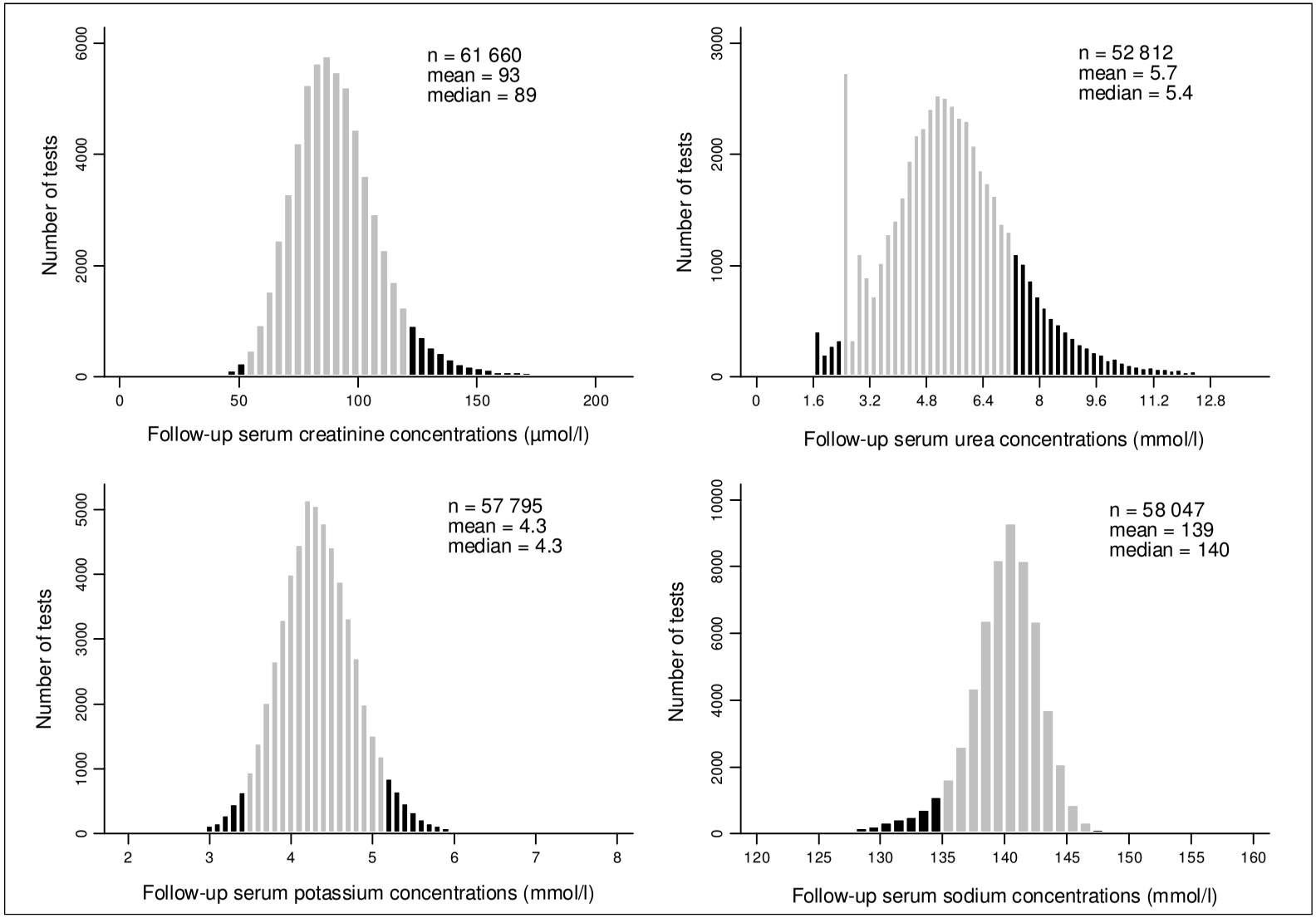


Figure 5.3 – Distribution of follow-up biochemical tests (black indicates concentrations outside standardized (95% CI) reference ranges)

The percentage of patients who had any biochemical monitoring in the first year of antihypertensive treatment increased from 37.5% in patients whose first prescription was in 2000 to 63.4% in patients whose first prescription was in 2003 (test for trend P<0.0005) (Figure 5.4).

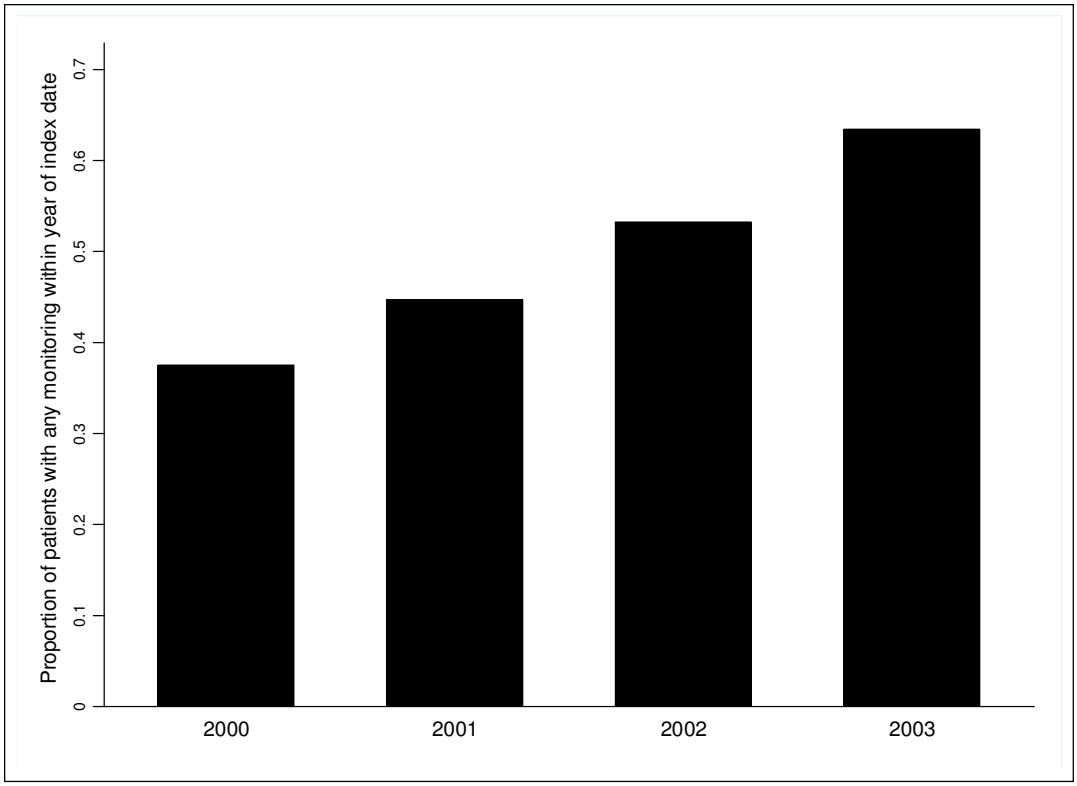


Figure 5.4 – Proportion of patients who had at least one biochemical monitoring test in the succeeding one year, by year of starting treatment

Over the time period examined, the proportion of patients within each GP practice with follow-up biochemical monitoring differed significantly (Figure 5.5). Some GP practices did not undertake any follow-up biochemical monitoring, while some practices monitored almost 9 in 10 of their patients treated with antihypertensive therapy.

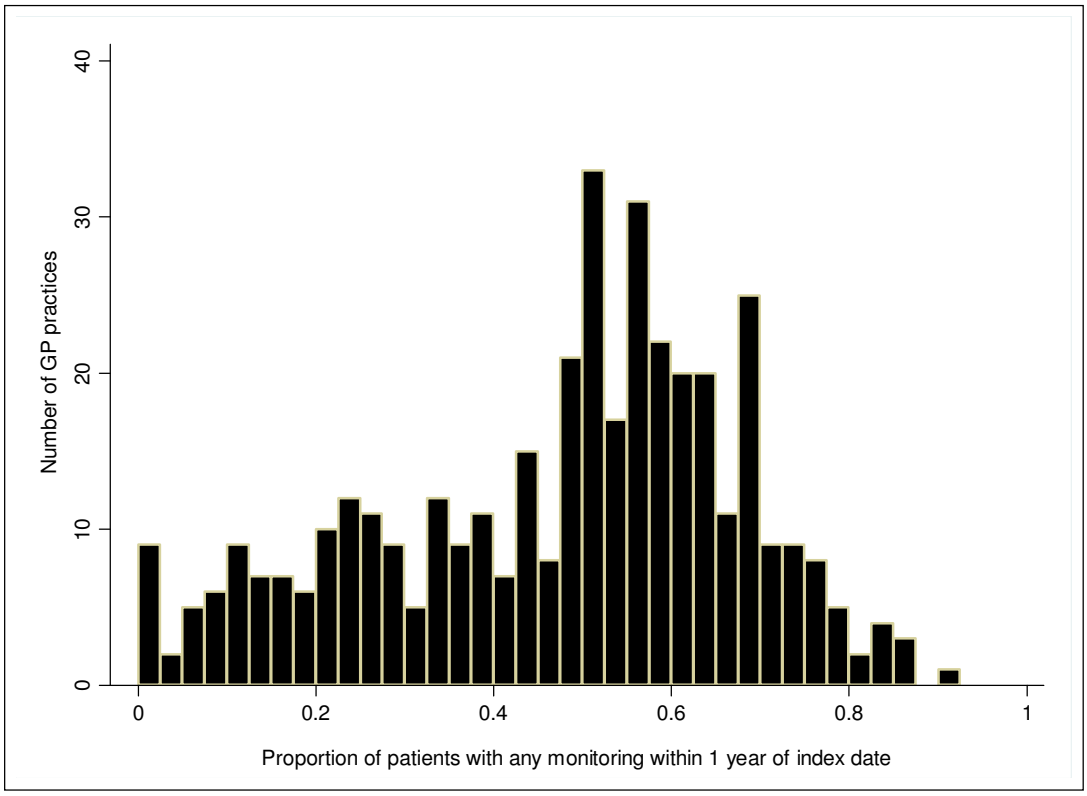


Figure 5.5 – Distribution of the proportion of patients with any monitoring within 1 year of index date by GP practice

The proportion of patients within a GP practice with follow-up biochemical monitoring increased significantly over time (Figure 5.6). In 2000, 13.6 % of practices undertook no biochemical monitoring of patients started antihypertensive therapy, compared with only 1.0% in 2003.

The mean number of repeat follow-up monitoring tests within one year in patients with at least one test was 1.68 (Table 5.6). The number of follow-up tests varied by drug class. Patients treated with ACE inhibitors and potassium-sparing or loop diuretics had a larger mean number of follow-up tests.

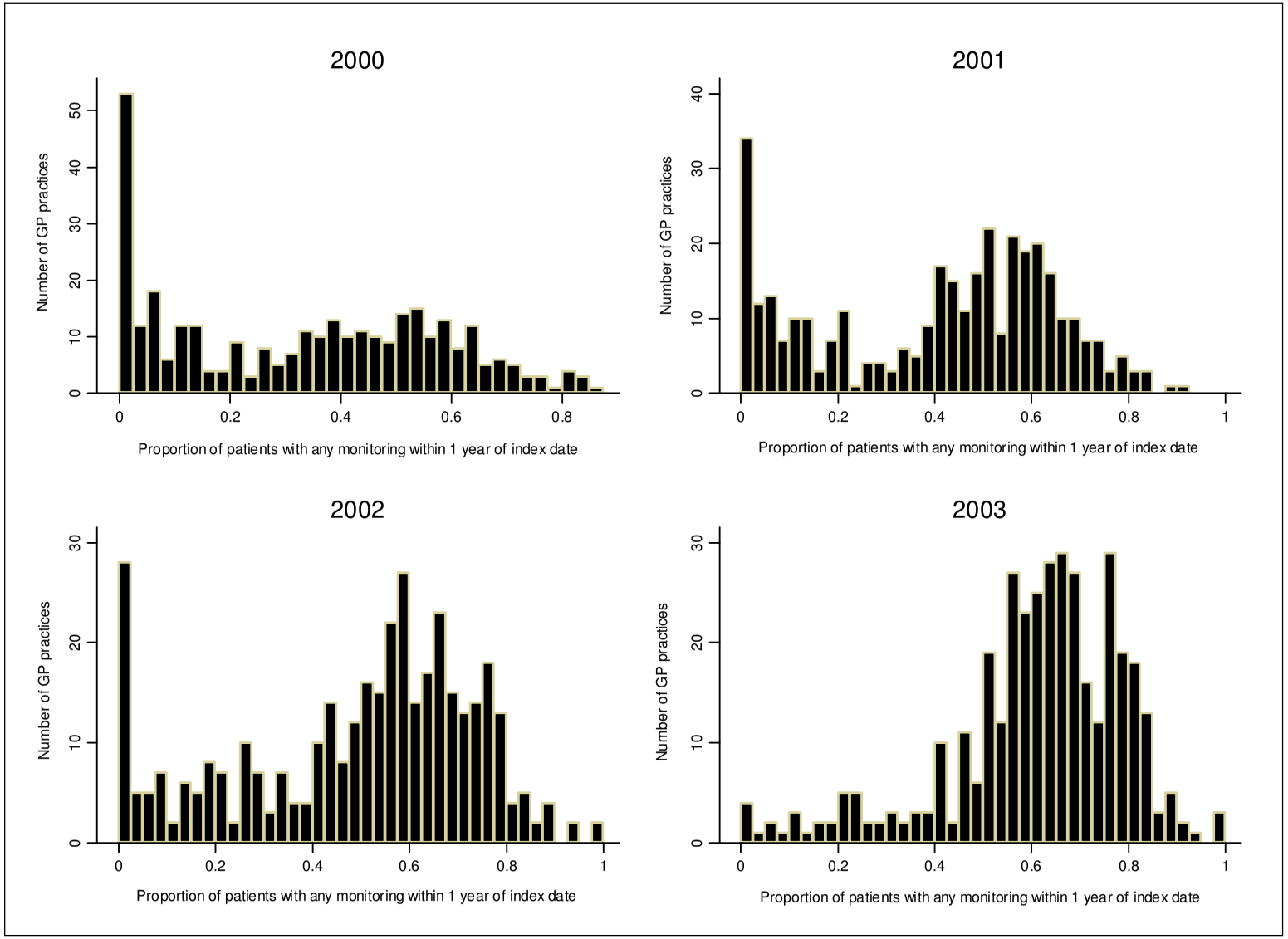


Figure 5.6 – Distribution of the proportion of patients with any monitoring within 1 year of index date by GP practice by year of initiation of antihypertensive treatment

Table 5.6 – Laboratory tests per person in the one year after first antihypertensive drug prescription by drug class (N=number of patients with at least one test)

	Creatinine			Urea			Sodium			Potassium		
Drug class	**N**	**Mean (median)**	**Range**	**N**	**Mean (median)**	**Range**	**N**	**Mean (median)**	**Range**	**N**	**Mean (median)**	**Range**
ACE inhibitor	7378	1.98 (1)	1–18	6086	1.95 (1)	1–18	6838	1.96 (1)	1–18	6848	1.94 (1)	1–18
Alpha-blocker	351	1.61 (1)	1–11	282	1.61 (1)	1–10	322	1.57 (1)	1–8	321	1.56 (1)	1–8
AT-II receptor antagonist	766	1.60 (1)	1–14	637	1.57 (1)	1–14	715	1.59 (1)	1–14	709	1.60 (1)	1–14
Beta-blocker	7686	1.57 (1)	1–18	6562	1.55 (1)	1–18	7369	1.56 (1)	1–18	7371	1.55 (1)	1–18
Ca-channel blocker	2555	1.67 (1)	1–15	2201	1.66 (1)	1–15	2407	1.65 (1)	1–15	2406	1.64 (1)	1–15
Combination	301	1.49 (1)	1–12	240	1.48 (1)	1–12	291	1.47 (1)	1–12	286	1.48 (1)	1–12
Loop or K- sparing diuretic	206	2.03 (1)	1–14	170	1.81 (1)	1–14	190	1.97 (1)	1–14	188	1.96 (1)	1–14
Thiazide diuretic	17315	1.62 (1)	1–30	15140	1.61 (1)	1–16	16463	1.62 (1)	1–30	16477	1.62 (1)	1–30
Total	**36558**	**1.68 (1)**	**1–30**	**31318**	**1.67 (1)**	**1–18**	**34595**	**1.68 (1)**	**1–30**	**34606**	**1.67**	**1–30**

A graph of time from index date to all of the biochemical follow-up monitoring tests reveals a sinusoidal pattern in the tail of the distribution, with peaks seen at seven-day intervals. All of the creatinine serum concentration measurements one month before and six months after initiation of antihypertensive therapy are presented in Figure 5.7 as an example. There was a large proportion of tests prior to the initiation of therapy and on days 0 to 2, indicating baseline testing. There also was a large initial peak of follow-up monitoring at day seven after the start of antihypertensive treatment and peaks are seen at seven-day intervals, which would suggest that monitoring tends to occurs at a certain number of weeks following the patient is started on therapy.

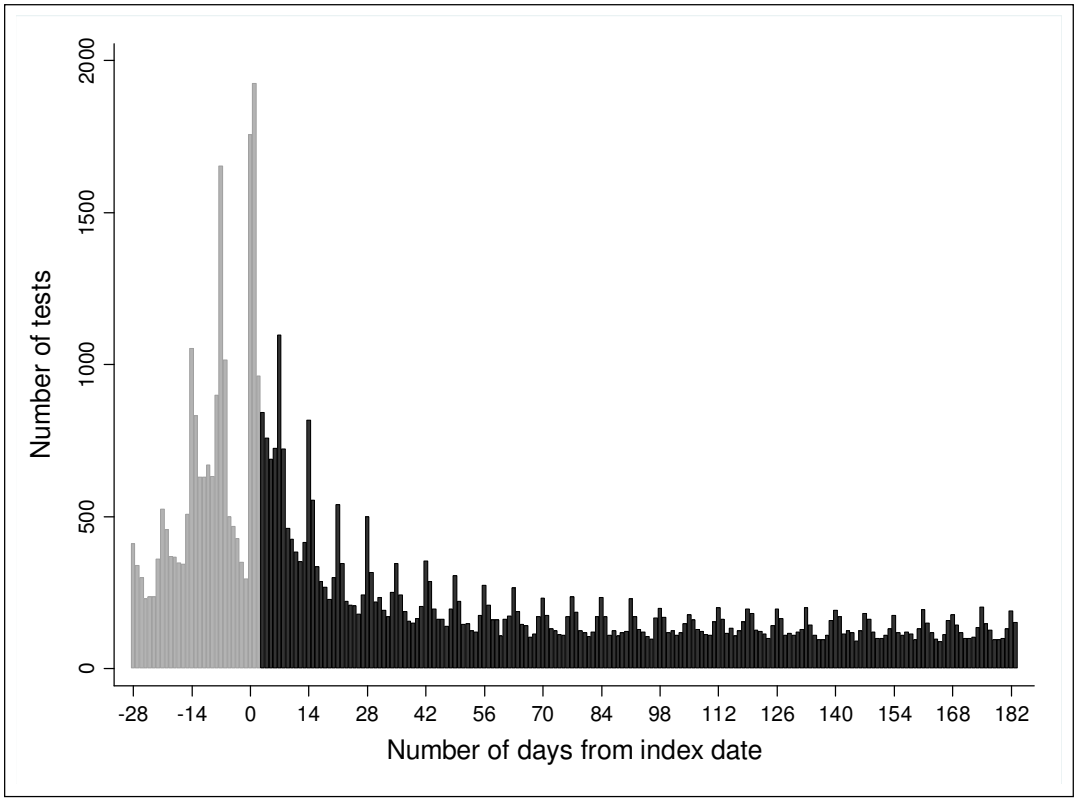


Figure 5.7 – Serum creatinine concentration tests in the one month before and six months after the index date (grey represents baseline testing)

The median interval between first antihypertensive prescription and first biochemical monitoring test within one year was 75 days, and the modal interval was seven days. The median interval between first antihypertensive prescription and first GP consultation within one year was 24 days, and the modal interval was 28 days (Figure 5.8).

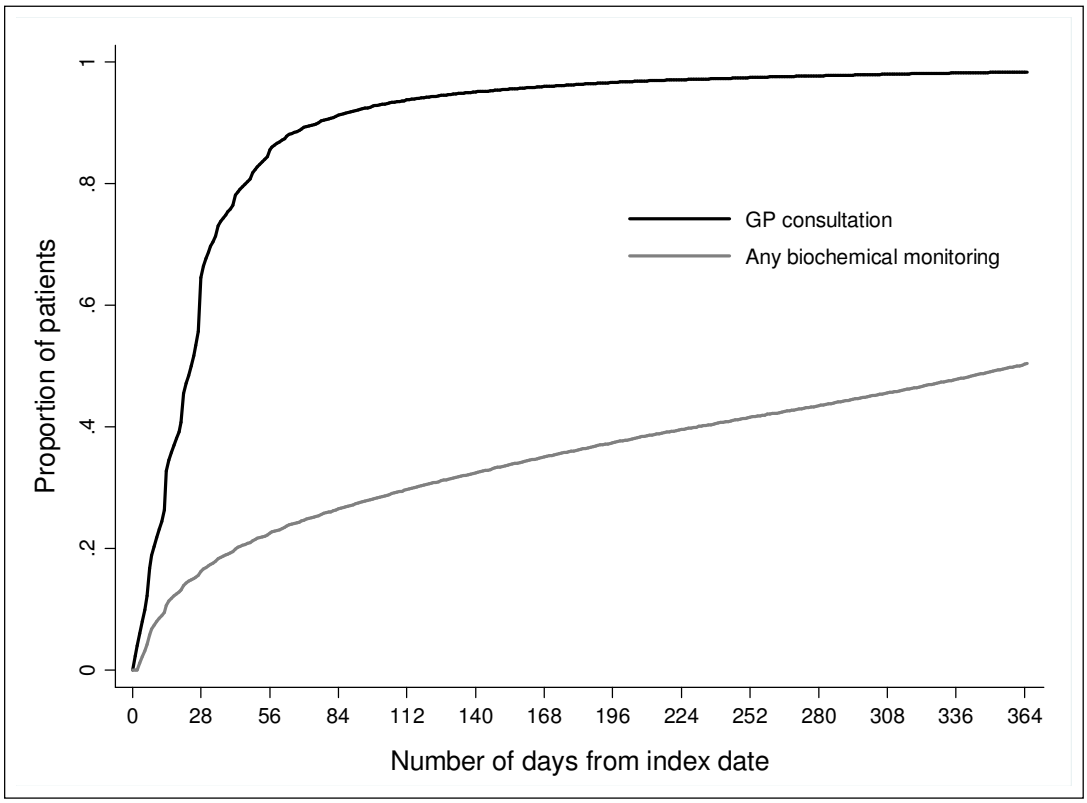


Figure 5.8 – Time to first biochemical monitoring test and first GP consultation in the twelve months after the first antihypertensive prescription

5.3.2 What is the regularity of follow-up monitoring?

In patients with any biochemical monitoring, the first differences varied considerably (Table 5.7). The time between tests was positively skewed with a mean value of 112 days and median value of 76 days. The second differences [see Figure 4.1 for a definition] between tests also varied widely and were positively skewed with a mean

value of 91 days and a median value of 63 days (Figure 5.9). Less than 2% of the patients with two or more tests had a mean second difference of 0 to 3 days, which would have suggested regular monitoring. Therefore very few patients had biochemical monitoring where the time intervals between tests were approximately equal.

Table 5.7 – First and second differences in the number of days between biochemical laboratory tests

	Creatinine	**Urea**	**Sodium**	**Potassium**
First differences (Δt)				
Mean (SD)	111.5 (102.6)	111.6 (103.1)	111.3 (102.9)	111.9 (103.1)
Median	76	76	76	76
Range	1–365	1–365	1–365	1–365
Interquartile range	26–178	26–180	25–179	26–180
Second differences (ΔΔt)				
Mean (SD)	90.8 (85.2)	91.5 (85.6)	91.1 (85.4)	91.1 (85.5)
Median	63	63	63	63
Range	0–358	0–358	0–358	0–358
Interquartile range	22–138	22–139	22–139	22–139
Number of patients with mean second difference between:				
0–3 days, n (%)	211/14659 (1.4)	178/12288 (1.4)	217/13776 (1.6)	215/13659 (1.6)
0–7 days, n (%)	474/14659 (3.2)	409/12288 (3.3)	471/13776 (3.4)	467/13659 (3.4)

Mean is the group mean and not the mean intra patient value, SD = standard deviation

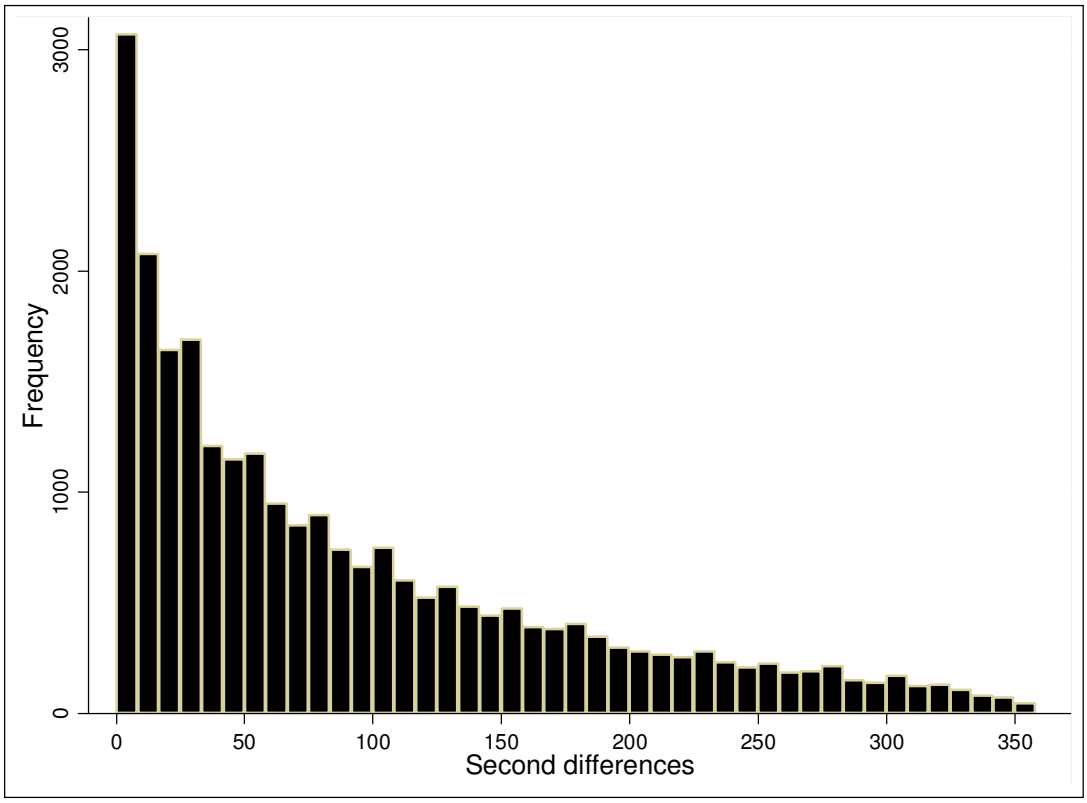


Figure 5.9 – Second differences in time between follow-up monitoring

5.3.3 How responsive is follow-up monitoring?

5.3.3.1 The responsiveness of follow-up monitoring to abnormal tests

When an abnormal serum concentration was identified, a record for a subsequent follow-up test was significantly more likely to be recorded (Table 5.8). However, the density of serum concentration tests was significantly lower after an abnormal test for all tests except for potassium (Table 5.9). Therefore when an abnormal laboratory test was recorded, a patient was significantly more likely to have a follow-up test, although not a significantly higher number of tests per month after the abnormal test.

Table 5.8 – Proportion of laboratory tests with a subsequent follow-up test

	Abnormal serum concentration with follow-up, n (%)	Normal serum concentration with follow-up, n (%)	P value
Creatinine	3037/5973 (50.9)	22065/55687 (39.6)	<0.0005
Urea	5606/12802 (43.8)	15258/39380 (38.8)	<0.0005
Sodium	2023/4091 (49.5)	21429/53956 (39.7)	<0.0005
Potassium	2608/5116 (51.0)	20581/52679 (39.1)	<0.0005

Table 5.9 – Density of serum concentration tests prior to and after an abnormal test result

	Density of tests (no. of tests per 28 days), mean (SD)		
	Prior to abnormal test	After abnormal test	P value†
Creatinine	0.189 (0.32)	0.157 (0.30)	<0.0005
Urea	0.157 (0.32)	0.122 (0.29)	<0.0005
Sodium	0.179 (0.33)	0.153 (0.53)	<0.0005
Potassium	0.170 (0.35)	0.158 (0.33)	0.61

† Wilcoxon matched paired signed-rank test

5.3.3.2 The responsiveness of follow-up monitoring to co-prescription of antihypertensive therapies

Very few patients had records of co-prescription of two drugs that could lead to potential ADRs. When spironolactone or an AT-II receptor antagonist was prescribed at the same time as an ACE inhibitor, there was a significant decrease in the density of serum potassium concentration tests after the date of co-prescription (Table 5.10). No statistically significant difference was observed in the number of tests prior to or after the concomitant treatment with a thiazide diuretic and a loop diuretic.

Table 5.10 – Density of serum concentration tests prior to and after the co-prescription of potentially interacting drugs.

Concomitant therapy	Potential ADR	Number of patients with concomitant therapy	Biochemical Test	Density of tests (no. of tests per 28 days), mean (SD)		
				Prior	After	P value†
ACE inhibitor & spironolactone	Hyperkalaemia	62	Potassium	0.197 (0.26)	0.147 (0.26)	0.0003
ACE inhibitor & AT-II receptor antagonist	Hyperkalaemia	2913	Potassium	0.185 (0.32)	0.088 (0.18)	<0.0005
Thiazide diuretic & loop diuretic	Hyperkalaemia	451	Potassium	0.113 (0.30)	0.172 (0.74)	0.25

† Wilcoxon matched paired signed-rank test

5.3.3.3 The responsiveness of follow-up monitoring to changes in treatment dose

Almost 40% (29550) of patients had a change in the dose of the first antihypertensive drug. 83% of the 56247 dose changes involved an increase in the prescribed daily dose. The density of tests was significantly smaller after a change in drug dose (Table 5.11).

Table 5.11 – Density of serum concentration tests prior to and after a dose change

	Density of tests (no. of tests per 28 days), mean (SD)		
	Prior to dose change	After dose change	P value†
Creatinine	0.173 (0.35)	0.089 (0.22)	<0.0005
Urea	0.169 (0.5)	0.086 (0.22)	<0.0005
Sodium	0.167 (0.35)	0.085 (0.22)	<0.0005
Potassium	0.167 (0.35)	0.085 (0.21)	<0.0005

† Wilcoxon matched paired signed-rank test

5.4 What proportion of patients had both baseline testing and follow-up monitoring?

In 17 445/74 096 patients (23.5%), both baseline testing and follow-up monitoring within one year were performed. Over a third of patients treated with an ACE inhibitor had both baseline testing and follow-up monitoring, compared with only 14% of patients started on combination antihypertensive treatment (Table 5.12).

Table 5.12 – Proportion of patients with records of both baseline and follow-up monitoring, by drug class

Drug class	No. of patients treated with drug	Creatinine n (%)	Sodium n (%)	Potassium n (%)
ACE inhibitor	11 245	4 375 (38.9)	3 925 (34.9)	3 892 (34.6)
Alpha-blocker	753	156 (20.7)	138 (18.3)	135 (17.9)
AT-II receptor antagonist	1 630	373 (22.9)	342 (21.0)	340 (20.9)
Beta-blocker	17 977	3 136 (17.4)	2 941 (16.4)	2 907 (16.2)
Ca-channel blocker	5 651	1 106 (19.6)	1 020 (18.1)	1 014 (17.9)
Combination	771	116 (15.1)	109 (14.1)	109 (14.1)
Loop or K-sparing diuretic	416	83 (20.0)	77 (18.5)	77 (18.5)
Thiazide diuretic	35 653	7 571 (21.2)	7 096 (19.9)	7 031 (19.7)

The proportion of patients with at least one abnormal result increased from 13% at baseline to 24% with treatment (Table 5.13). Of the 15 215 patients with normal results at baseline, 2883 (18.9%) developed an abnormal test within one year of starting treatment (Figure 5.10).

Table 5.13 – Proportion of patients with an abnormal biochemical test in those patients with both baseline testing and follow-up monitoring within one year

	Any test (N=17 445)	Creatinine (N=16 916)	Sodium (N=15 648)	Potassium (N=15 505)
Abnormal baseline serum concentration, n (%)	2 230 (12.8)	803 (4.8)	584 (3.7)	908 (5.9)
Any abnormal serum concentration within one year of treatment, n (%)	4 188 (24.0)	1 302 (7.7)	1 324 (8.5)	2 039 (13.2)

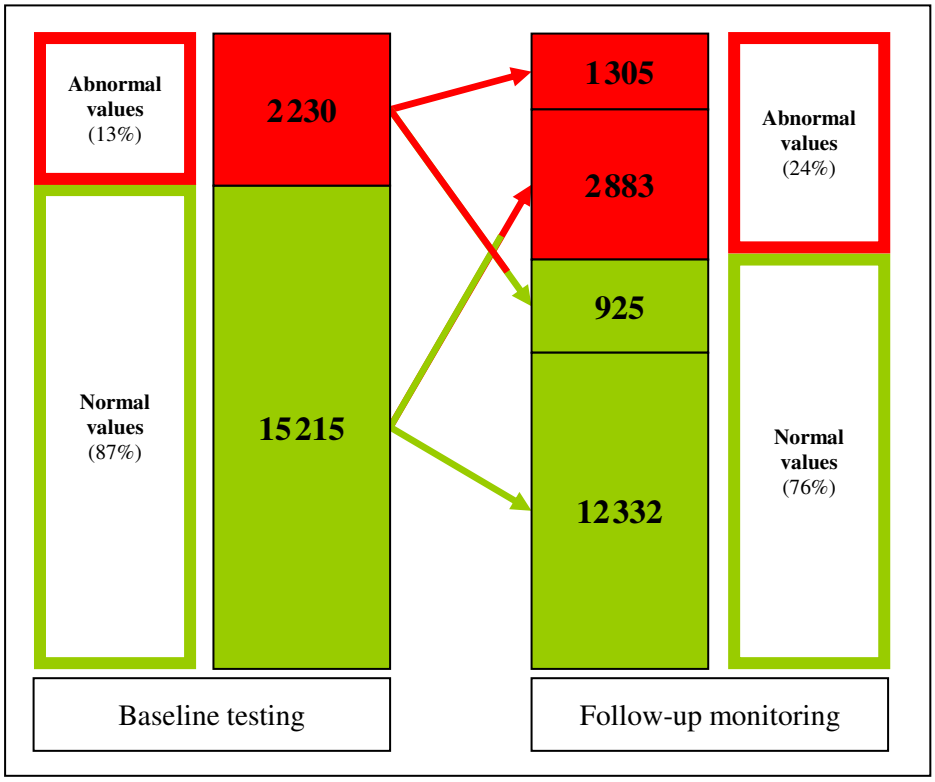


Figure 5.10 – Distribution of abnormal tests in patients who had both baseline testing and follow-up monitoring

5.5 What are the patient factors associated with biochemical monitoring?

The patient factors associated with monitoring within six months and one year are shown in Table 5.14 and Table 5.15. Patients with pre-existing diabetes, patients started on ACE inhibitors at baseline, and those with a first antihypertensive prescription in 2003 were more likely to be monitored. In the multivariable model, monitoring was again significantly associated the presence of diabetes, first-line antihypertensive therapy with ACE inhibitors, and year of first prescription. Monitoring within one year was also significantly associated with increasing age, baseline testing, and with socio-economic status, although the magnitude of these associations was slight.

Table 5.14 – Logistic regression model of any monitoring in six months following first prescription (n is the number of patients who exhibit the characteristic from N, the total population assessed for this characteristic)

Baseline patient characteristic	n/N	Unadjusted OR 95% CI		P value	Adjusted OR* 95% CI		P value
Sex							
Female	13 768/38 751	1			1		
Male	13 178/35 345	1.08	1.05–1.11	<0.0005	1.00	0.97–1.04	0.91
Age							
<40	1 112/3 492	1			1		
40–49	3 599/10 388	1.12	1.03–1.22	0.007	1.09	0.99–1.19	0.09
50–59	7 118/20 360	1.14	1.05–1.23	0.001	1.12	1.03–1.22	0.009
60–69	7 214/19 312	1.26	1.17–1.36	<0.0005	1.24	1.14–1.36	<0.0005
70–79	5 830/14 971	1.35	1.25–1.46	<0.0005	1.35	1.23–1.48	<0.0005
80–89	1 936/5 190	1.26	1.15–1.38	<0.0005	1.26	1.13–1.41	<0.0005
90–100	127/383	1.05	0.84–1.31	0.683	1.17	0.84–1.63	0.34
Smoking status							
Never	13 710/38 166	1			1		
Ever	12 537/33 118	1.09	1.05–1.12	<0.0005	1.08	1.04–1.12	<0.0005
Diabetes status							
No	23 271/67 659	1			1		

Baseline patient characteristic	n/N	Unadjusted OR 95% CI		P value	Adjusted OR* 95% CI		P value
Yes	3675/6437	2.54	2.41–2.67	<0.0005	2.03	1.91–2.16	<0.0005
SES quintile (0=least deprived)							
0	5823/15660	1			1		
1	5011/13205	1.03	0.98–1.08	0.182	1.06	1.00–1.12	0.04
2	5318/15005	0.93	0.89–0.97	0.002	0.92	0.87–0.97	0.002
3	4760/14154	0.86	0.82–0.90	<0.0005	0.88	0.83–0.93	<0.0005
4	6034/16072	1.02	0.97–1.06	0.508	1.06	1.00–1.12	0.02
Hypertension							
Normal or mild	4147/11373	1			1		
Moderate	10836/30241	0.97	0.93–1.02	0.232	1.07	1.02–1.13	0.005
Severe	11098/28529	1.11	1.06–1.16	<0.0005	1.31	1.25–1.38	<0.0005
BMI							
Underweight	243/654	1.01	0.86–1.19	0.860	1.01	0.85–1.21	0.88
Normal	5905/16039	1			1		
Overweight	9334/25392	0.98	0.96–1.03	0.907	0.98	0.94–1.03	0.42
Obese	7840/20606	1.05	1.00–1.10	0.016	1.02	0.97–1.07	0.37
Drug Therapy							
Thiazide diuretic	12492/35653	1			1		
ACE inhibitor	6049/11245	2.16	2.07–2.25	<0.0005	1.93	1.82–2.03	<0.0005
Alpha-blocker	220/753	0.93	0.77–0.90	0.001	0.76	0.63–0.91	0.003
AT-II receptor antagonist	576/1630	1.01	0.91–1.12	0.804	0.99	0.87–1.11	0.81
Beta-blocker	5441/17977	0.80	0.77–0.84	<0.0005	0.87	0.83–0.91	<0.0005
Ca-channel blocker	1808/5651	0.87	0.82–0.93	<0.0005	0.89	0.83–0.96	0.001
Loop or K-sparing diuretic	149/416	1.03	0.85–1.27	0.740	1.41	1.10–1.81	0.006
Combination preparation	211/771	0.70	0.60–0.82	<0.0005	0.79	0.66–0.95	0.01
Any baseline testing							
No	15419/43002				1		
Yes	11527/31094	1.05	1.02–1.09	0.001	0.80	0.77–0.83	<0.0005
Year of first prescription							
2000	4251/16508	1			1		
2001	5742/17927	1.36	1.30–1.42	<0.0005	1.35	1.28–1.43	<0.0005
2002	7543/19532	1.81	1.73–1.90	<0.0005	1.77	1.68–1.87	<0.0005
2003	9410/20129	2.53	2.42–2.64	<0.0005	2.56	2.43–2.69	<0.0005

* adjusted for sex, age, smoking status, diabetes, socio-economic status (SES), BMI, hypertension, drug therapy, any baseline testing, and year of first prescription

Table 5.15 – Logistic regression model of any monitoring in one year following first prescription (n is the number of patients who exhibit the characteristic from N, the total population assessed for this characteristic)

Baseline patient characteristic	n/N	Unadjusted OR 95% CI		P value	Adjusted OR* 95% CI		P value
Sex							
Female	19 151/38 751	1			1		
Male	18 214/35 345	1.09	1.06–1.12	<0.0005	1.02	0.98–1.05	0.32
Age							
<40	1 486/3 492	1			1		
40–49	4 882/10 388	1.19	1.11–1.29	<0.0005	1.19	1.09–1.30	<0.0005
50–59	9 811/20 360	1.26	1.17–1.35	<0.0005	1.27	1.17–1.38	<0.0005
60–69	10 122/19 312	1.49	1.38–1.60	<0.0005	1.52	1.39–1.65	<0.0005
70–79	8 180/14 971	1.63	1.51–1.75	<0.0005	1.71	1.56–1.87	<0.0005
80–89	2 701/5 190	1.46	1.34–1.60	<0.0005	1.57	1.41–1.75	<0.0005
90–100	183/383	1.24	1.00–1.53	0.050	1.52	1.10–2.11	0.01
Smoking status							
Never	19 149/38 166	1			1		
Ever	17 274/33 118	1.08	1.05–1.12	<0.0005	1.06	1.02–1.10	0.001
Diabetes status							
No	32 491/67 659	1			1		
Yes	4 874/6 437	3.38	3.18–3.58	<0.0005	2.77	2.58–2.97	<0.0005
SES quintile (0=least deprived)							
0	7 843/15 660	1			1		
1	6 891/13 205	1.09	1.04–1.14	<0.0005	1.11	1.05–1.17	<0.0005
2	7 377/15 005	0.96	0.92–1.00	0.107	0.95	0.90–1.00	0.06
3	6 731/14 154	0.90	0.86–0.95	<0.0005	0.94	0.89–0.99	0.02
4	8 523/16 072	1.13	1.08–1.18	<0.0005	1.17	1.11–1.24	<0.0005
Hypertension							
Normal or mild	5 829/11 373	1			1		
Moderate	15 334/30 241	0.98	0.94–1.02	0.320	1.07	1.02–1.12	0.007
Severe	14 975/28 529	1.05	1.00–1.10	0.026	1.22	1.16–1.28	<0.0005
BMI							
Underweight	342/654	1.04	0.89–1.22	0.586	1.02	0.86–1.21	0.81
Normal	8 213/16 039	1			1		
Overweight	13 125/25 392	1.02	0.98–1.06	0.338	1.01	0.97–1.05	0.67
Obese	10 817/20 606	1.05	1.01–1.10	0.014	1.05	1.00–1.10	0.03
Drug Therapy							
Thiazide diuretic	17 706/35 653	1			1		
ACE inhibitor	7 473/11 245	2.00	1.92–2.10	<0.0005	1.62	1.53–1.71	<0.0005

Baseline patient characteristic	n/N	Unadjusted OR 95% CI		P value	Adjusted OR* 95% CI		P value
Alpha-blocker	360/753	0.93	0.80–1.07	0.314	0.93	0.79–1.11	0.43
AT-II receptor antagonist	783/1 630	0.94	0.85–1.03	0.199	0.89	0.79–0.99	0.05
Beta-blocker	7 874/17 977	0.78	0.76–0.82	<0.0005	0.89	0.85–0.92	<0.0005
Ca-channel blocker	2 643/5 651	0.89	0.84–0.94	<0.0005	0.91	0.85–0.97	0.005
Loop or K-sparing diuretic	213/416	1.05	0.87–1.28	0.598	1.32	1.02–1.69	0.03
Combination preparation	314/771	0.70	0.60–0.81	<0.0005	0.83	0.70–0.99	0.03
Any baseline testing							
No	19 920/43 002				1		
Yes	17 445/31 094	1.48	1.44–1.53	<0.0005	1.15	1.11–1.19	<0.0005
Year of first prescription							
2000	6 189/16 508	1			1		
2001	8 008/17 927	1.34	1.29–1.41	<0.0005	1.31	1.25–1.38	<0.0005
2002	10 401/19 532	1.90	1.82–1.98	<0.0005	1.83	1.74–1.92	<0.0005
2003	12 767/20 129	2.89	2.77–3.01	<0.0005	2.80	2.66–2.94	<0.0005

* adjusted for sex, age, smoking status, diabetes, socio-economic status (SES), BMI, hypertension, drug therapy, any baseline testing, and year of first prescription

5.6 What is the association between follow-up monitoring and adverse patient outcomes?

Overall, the rate of biochemical ADRs within six months and one year of starting treatment was low and varied by drug class (Table 5.16). Patients treated with ACE inhibitors or thiazide diuretics had a higher rate of hyponatraemia. Hyperkalaemia was more frequently detected in patients treated with ACE inhibitors and hypokalaemia was most often seen in patients treated with thiazide diuretics.

Table 5.16 – Frequency of biochemical ADRs within six months and one year of the initiation of antihypertensive treatment*

First drug class initiated	Number of patients started on drug	Hyponatraemia		Hyperkalaemia		Hypokalaemia	
		Within 6 months	Within 1 year	Within 6 months	Within 1 year	Within 6 months	Within 1 year
ACE inhibitor, n (%)	11 245	290 (2.6; 5.2)	492 (4.4; 7.0)	212 (1.9; 3.8)	308 (2.7; 4.5)	79 (0.7; 1.4)	125 (1.1; 1.8)
Alpha blocker, n (%)	753	8 (1.1; 4.1)	19 (2.5; 5.9)	3 (0.4; 1.5)	6 (0.8; 1.9)	4 (0.5; 2.1)	13 (1.7; 4.0)
AT-II receptor antagonist, n (%)	1 630	19 (1.2; 3.6)	29 (1.8; 4.1)	9 (0.6; 1.7)	20 (1.2; 2.8)	4 (0.3; 0.8)	9 (0.6; 1.3)
Beta-blocker, n (%)	17 977	201 (1.1; 4.0)	375 (2.1; 5.1)	167 (0.9; 3.3)	261 (1.5; 3.5)	103 (0.6; 2.0)	189 (1.1; 2.6)
Ca-channel blocker, n (%)	5 651	64 (1.1; 3.9)	128 (2.3; 5.3)	44 (0.8; 2.7)	67 (1.2; 2.8)	38 (0.7; 2.3)	90 (1.6; 3.7)
Loop or K-sparing diuretic, n (%)	416	10 (2.4; 7.5)	19 (4.6; 10.0)	1 (0.1; 0.8)	10 (2.4; 5.3)	11 (1.4; 5.5)	5 (1.2; 2.7)
Combination preparation, n (%)	771	14 (1.8; 7.3)	21 (2.7; 7.2)	6 (1.4; 3.2)	3 (0.4; 1.0)	2 (0.5; 1.1)	20 (2.6; 7.0)
Thiazide diuretic, n (%)	35 653	904 (2.5; 7.8)	1 525 (4.3; 9.3)	162 (0.5; 1.4)	285 (0.8; 1.7)	691 (1.9; 6.0)	1 080 (3.0; 6.6)
Total, n (%)	**74 096**	**1510 (2.0; 6.1)**	**2 608 (3.5; 7.5)**	**604 (0.8; 2.4)**	**960 (1.3; 2.8)**	**932 (1.3; 3.8)**	**1 531 (2.1; 4.4)**

* presented as a percentage of patients started on the drug and a percentage of patients with biochemical monitoring

657 patients (0.9%) died within one year of beginning antihypertensive treatment. The death rate was significantly higher in those who had any biochemical monitoring than in those patients who did not. This increase was not explained by confounding as adjustment for potential confounding factors did not reduce the magnitude of association (adjusted hazard ratio 1.37, 95% CI 1.12–1.67) (Table 5.17).

2499 patients (3.4%) were admitted to hospital at least once within one year of starting antihypertensive treatment. 456 patients had more than one admission (range 1–7). Patients were more likely to be admitted to hospital if their biochemistry had been monitored after beginning treatment, even when allowance was made for sex, age, smoking, presence of diabetes, socio-economic status, and first-line drug therapy (adjusted hazard ratio 1.27, 95% CI 1.16–1.39) (Table 5.18).

Table 5.17 – Hazard ratios (with 95% confidence intervals) for death in the one year following first antihypertensive prescription

	Unadjusted HR		**P value**	**Adjusted HR***		**P value**
	HR	**95% CI**		**HR**	**95% CI**	
Any monitoring on or before event						
No	1			1		
Yes	1.28	1.09–1.50	0.003	1.37	1.12–1.67	0.002
Sex						
Female	1			1		
Male	1.26	1.08–1.47	0.003	1.51	1.24–1.85	<0.0005
Age						
<50	1			1		
50–59	3.02	1.81–5.04	<0.0005	3.69	1.94–7.01	<0.0005
60–69	5.21	3.19–8.54	<0.0005	5.83	3.12–10.9	<0.0005
70–79	11.8	7.30–19.1	<0.0005	12.0	6.49–22.3	<0.0005
80–89	24.9	15.3–40.6	<0.0005	29.1	15.6–54.5	<0.0005
90–100	69.7	39.3–123	<0.0005	71.9	32.4–159	<0.0005
Smoking						
Never	1			1		
Ever	2.15	1.76–2.62	<0.0005	2.15	1.75–2.64	<0.0005

	Unadjusted HR		P value	Adjusted HR*		P value
	HR	**95% CI**		**HR**	**95% CI**	
Diabetes status						
No	1			1		
Yes	1.59	1.26–1.99	<0.0005	1.82	1.35–2.45	<0.0005
SES quintile						
0 (least deprived)	1			1		
1	1.35	1.05–1.72	0.019	1.80	1.31–2.47	<0.0005
2	1.18	0.92–1.51	0.199	1.48	1.07–2.04	0.02
3	1.17	0.91–1.51	0.217	1.44	1.04–1.99	0.03
4	1.21	0.95–1.55	0.116	1.45	1.05–1.99	0.02
Hypertension						
Normal or mild	1			-		
Moderate	0.82	0.65–1.05	0.118	-		
Severe	1.22	0.97–1.54	0.089	-		
Drug therapy						
Thiazide diuretic	1			1		
ACE inhibitor	0.89	0.70–1.13	0.345	0.87	0.63–1.20	0.38
Alpha-blocker	1.22	0.61–2.47	0.573	0.37	0.09–1.50	0.16
AT-II receptor antagonist	0.63	0.33–1.23	0.179	0.27	0.07–1.10	0.07
Beta-blocker	0.74	0.60–0.92	0.006	1.24	0.96–1.60	0.10
Ca-channel blocker	1.54	1.19–1.98	0.001	1.34	0.98–1.84	0.07
Loop/K-sparing diuretic	9.63	6.72–13.8	<0.0005	4.40	2.64–7.34	<0.0005
Combination preparation	2.88	1.81–4.58	<0.0005	2.57	1.36–4.84	0.004

* adjusted for any monitoring, age, sex, smoking status, diabetes, SES, and first antihypertensive therapy

Table 5.18 – Hazard ratios (with 95% confidence intervals) for first hospital admission in the one year following antihypertensive prescription

	Unadjusted HR		P value	Adjusted HR*		P value
	HR	95% CI		HR	95% CI	
Any monitoring on or before event						
No	1					
Yes	1.30	1.20–1.42	<0.0005	1.27	1.16–1.39	<0.0005
Sex						
Female				1		
Male	1.03	0.96–1.12	0.391	1.06	0.97–1.15	0.21
Age						
<50	1			1		
50–59	1.00	0.96–1.15	0.957	1.05	0.91–1.21	0.50
60–69	1.33	1.17–1.51	<0.0005	1.33	1.16–1.52	<0.0005
70–79	1.71	1.50–1.95	<0.0005	1.77	1.54–2.03	<0.0005
80–89	2.15	1.84–2.51	<0.0005	2.12	1.78–2.52	<0.0005
90–100	2.14	1.36–3.36	0.001	2.11	1.23–3.62	0.01
Smoking						
Never	1			1		
Ever	1.14	1.05–1.23	0.002	1.12	1.03–1.22	0.01
Diabetes status						
No	1			1		
Yes	1.17	1.03–1.34	0.017	1.11	0.96–1.18	0.17
SES quintile						
0 (least deprived)	1			1		
1	0.95	0.83–1.10	0.520	0.95	0.82–1.11	0.55
2	1.19	1.04–1.36	0.010	1.20	1.05–1.38	0.01
3	1.79	1.58–2.02	<0.0005	1.82	1.61–2.08	<0.0005
4	1.42	1.25–1.61	<0.0005	1.38	1.21–1.58	<0.0005
Hypertension						
Normal or mild	1			1		
Moderate	0.87	0.78–0.98	0.024	0.86	0.76–0.97	0.02
Severe	1.07	0.95–1.20	0.272	0.95	0.84–1.08	0.44
Drug therapy						
Thiazide diuretic	1			1		
ACE inhibitor	1.03	0.91–1.16	0.620	1.03	0.90–1.18	0.67
Alpha-blocker	1.29	0.90–1.86	0.163	1.00	0.65–1.54	0.99
AT-II receptor antagonist	1.01	0.76–1.34	0.941	1.04	0.77–1.41	0.80
Beta-blocker	1.12	1.01–1.23	0.209	1.25	1.12–1.39	<0.0005
Ca-channel blocker	1.63	1.43–1.85	<0.0005	1.52	1.32–1.75	<0.0005
Loop/K-sparing diuretic	2.94	2.10–4.11	<0.0005	2.02	1.35–3.05	0.001
Combination preparation	0.98	0.64–1.48	0.922	0.85	0.52–1.37	0.50

* adjusted for any monitoring, age, sex, smoking status, diabetes, SES, hypertension and first antihypertensive therapy

Over half (52.7%) of patients discontinued the antihypertensive drug initiated on the index date within one year. The mean time to discontinuation was 120 days (median 86). Of the 39019 patients who discontinued treatment, 7316 (18.7%) had a discontinuation date of 28 days. This would suggest that a significant proportion of patients discontinued their antihypertensive treatment after one course of therapy. Any biochemical monitoring was significantly associated with a small increase in the risk of discontinuation of the first antihypertensive prescription (adjusted HR 1.05, 95% CI 1.03–1.08) (Table 5.19).

Table 5.19 – Hazard ratios (with 95% confidence intervals) for discontinuation of the first antihypertensive prescription in the one year following the start of treatment

	Unadjusted HR		P value	Adjusted HR*		P value
	HR	95% CI		HR	95% CI	
Any monitoring on or before event						
No	1			1		
Yes	1.02	0.99–1.04	0.087	1.05	1.03–1.08	<0.0005
Sex						
Female	1			1		
Male	1.00	0.98–1.02	0.687	1.01	0.99–1.03	0.300
Age						
<50	1			1		
50–59	0.83	0.81–0.85	<0.0005	0.83	0.80–0.86	<0.0005
60–69	0.77	0.75–0.79	<0.0005	0.76	0.73–0.78	<0.0005
70–79	0.80	0.77–0.82	<0.0005	0.78	0.75–0.81	<0.0005
80–89	0.94	0.90–0.98	<0.0005	0.89	0.85–0.94	<0.0005
90–100	1.08	0.95–1.22	0.262	1.03	0.88–1.20	0.71
Smoking						
Never	1			1		
Ever	1.02	0.99–1.04	0.057	1.02	1.00–1.04	0.05
Diabetes status						
No	1			1		
Yes	0.86	0.83–0.89	<0.0005	0.89	0.85–0.94	<0.0005

	Unadjusted HR		P value	Adjusted HR*		P value
	HR	95% CI		HR	95% CI	
SES quintile						
0 (least deprived)	1			1		
1	1.02	0.99–1.05	0.244	0.99	0.96–1.03	0.82
2	1.02	0.99–1.06	0.136	1.02	0.99–1.05	0.26
3	1.10	1.07–1.14	<0.0005	1.08	1.05–1.12	<0.0005
4	1.13	1.09–1.16	<0.0005	1.10	1.06–1.13	<0.0005
Hypertension						
Normal or mild	1			1		
Moderate	0.92	0.89–0.95	<0.0005	0.92	0.89–0.95	<0.0005
Severe	0.93	0.90–0.96	<0.0005	0.93	0.90–0.96	<0.0005
Drug therapy						
Thiazide diuretic	1			1		
ACE inhibitor	0.91	0.88–0.93	<0.0005	0.90	0.87–0.93	<0.0005
Alpha-blocker	1.00	0.91–1.11	0.883	1.00	0.90–1.12	0.94
AT-II receptor	0.83	0.77–0.90	<0.0005	0.80	0.74–0.86	<0.0005
Beta-blocker	0.97	0.95–0.99	0.016	0.94	0.92–0.97	<0.0005
Ca-channel blocker	1.11	1.07–1.15	<0.0005	1.10	1.06–1.15	<0.0005
Loop/K-sparing diuretic	2.30	2.05–2.57	<0.0005	2.30	2.03–2.60	<0.0005
Combination preparation	1.11	1.00–1.22	0.031	1.06	0.96–1.18	0.27

* adjusted for any monitoring, age, sex, smoking status, diabetes, SES, hypertension and first antihypertensive therapy

I defined drug discontinuation as no subsequent prescription 30 days after the expiry of the prescription. The delay of 30 days was to allow for alternative supply and less than perfect dosing compliance (section 3.3.6.1). The use of an alternative 60-day definition of discontinuation decreased the proportion of patients discontinuing the first-ever antihypertensive therapy within one year to 40.4% (Table 5.20). The use of a 90-day definition further decreased the proportion of patients who discontinued their first course of treatment to 35.7%. The predictive relationships of other patient factors in the multivariable models remained similar when the different discontinuation definitions were examined.

Table 5.20 – Sensitivity analysis comparing different discontinuation definitions

Discontinuation definition	Patients discontinuing treatment within 1 year, n (%)	Any monitoring on or before the day of discontinuation of the first antihypertensive prescription within 1 year					
		Unadjusted HR (95% CI)		P value	Adjusted HR* (95% CI)		P value
30-day definition	39019 (52.7)	1.02	0.99–1.04	0.087	1.05	1.03–1.08	<0.0005
60-day definition	29909 (40.4)	1.05	1.02–1.08	<0.0005	1.09	1.06–1.12	<0.0005
90-day definition	26467 (35.7)	1.07	1.04–1.10	<0.0005	1.11	1.09–1.15	<0.0005

* adjusted for any monitoring, age, sex, smoking status, diabetes, SES, hypertension and first antihypertensive therapy

Within the cohort of patients who had a serum concentration test within one year, a serum concentration test indicative of a biochemical ADR was significantly associated with an increased risk of death, hospital admission, and discontinuation of first antihypertensive treatment (Table 5.21).

Table 5.21 – Hazard ratios (with 95% confidence intervals) for adverse patient outcomes after a serum concentration test indicating a biochemical ADR

	Death	Hospital admission	Discontinuation of first antihypertensive treatment
	Adjusted HR* (95% CI)	Adjusted HR* (95% CI)	Adjusted HR* (95% CI)
Hyponatraemia	5.71 (3.61–9.04)	2.16 (1.53–3.03)	2.24 (1.99–2.52)
Hyperkalaemia	2.93 (1.65–5.18)	1.77 (1.26–2.50)	1.24 (1.09–1.40)
Hypokalaemia	4.34 (2.63–7.16)	1.54 (1.12–2.13)	1.39 (1.26–1.52)

* adjusted for age, sex, smoking status, diabetes, SES, hypertension and first antihypertensive therapy

5.7 Assessing the impact of the definition of baseline testing and follow-up monitoring

Baseline testing was defined as any record of a serum concentration in the six months before and up to and including two days after the start of treatment, while follow-up monitoring was defined as any record of a serum concentration from three days after the start of treatment. The decision was made to use this definition primarily because of potential delays in the uploading of laboratory test results as prior to 2003 the majority of laboratory data entry was not automated. This decision is supported by Figure 5.7, which demonstrates a large number of tests on days 0 and 1, followed by a significant drop in the number of tests.

If the definition of baseline testing is changed to include the first seven days after the initiation of antihypertensive therapy, the number of patients with baseline testing is increased from 31 094 (42%) to 35 388 (48%). The corollary of the change in definition is that the number of patients with follow-up monitoring within one year decreases from 37 365 (50%) to 33 071 (45%). Almost 14% (576) of the 4294 patients with a biochemical test on days three to seven had a record of baseline testing. Changing the definitions of baseline testing and follow-up monitoring would also increase the number of patients who had both baseline testing and follow-up monitoring from 17 445 (23.5%) to 19 550 (26.4%). Certainly there is a difference depending on the definition used, however this change is not very large and does not greatly affect the results.

5.8 Investigating the impact of immortal time bias on the results

5.8.1 Quantifying the impact of immortal time bias

The estimates for the association between biochemical monitoring and adverse patient outcomes presented in previous sections were adjusted for immortal time bias. This bias can arise when immortal time is misclassified and can create an artificial protective effect for an exposure (described previously in section 4.3.3).

I wished to illustrate the nature of immortal time bias through the use of a Poisson regression model. The rate ratios of the various adverse events associated with biochemical monitoring were estimated using the number of events and person-time. When data were unadjusted for immortal time, a decreased mortality rate with biochemical monitoring was observed (crude rate ratio 0.67, 95% CI 0.66–0.69) (Table 5.22). I also used a time-fixed Cox proportional hazards model where monitoring was not treated as a time-dependent covariate and immortal time was not correctly classified. A similar adjusted hazard ratio to the crude rate ratio was obtained (0.67, 95% CI 0.55–0.81). Therefore, this analysis, which did not take into account immortal time, suggested that biochemical monitoring had a protective effect.

There were 12013 person-years of follow-up during which patients were yet to have biochemical monitoring, which should have been ascribed to the unmonitored group. This accounted for approximately 32% of total follow-up time allocated to patients with monitoring. This immortal time, when correctly classified, resulted in a corrected crude rate ratio of 1.33 (95% CI 1.30–1.36), suggesting an increased risk of death in patients with biochemical monitoring. This result is similar to the adjusted hazard

ratio from the time-to-event analyses presented previously (1.37, 95% CI 1.12–1.67). Similar differences in results were obtained when the failure to control for immortal time bias in the analysis of biochemical monitoring and hospital admission and drug discontinuation was examined.

Table 5.22 – Rate ratios examining the relationships between biochemical monitoring and adverse patient outcomes using two approaches (one where immortal time bias is not adjusted for and another where it is)

	Number of patients	Number of events	Person-years	Rate (per 1000 per year)	Crude rate ratio (95% CI)	Adjusted Hazard Ratio†
Death						
Misclassified approach						
No monitoring within 1 year	36731	388	36033	10.8	1.0	1.0
Monitoring within 1 year	37365	269	37070	7.3	0.67 (0.66–0.69)	0.67 (0.55–0.81)
Correctly classified approach						
No monitoring within 1 year	74096	388	48045	8.1	1.0	1.0
No monitoring within 1 year	36731	388	36033			
From study entry until monitoring	37365	0	12013			
Monitoring within 1 year	37365	269	25057	10.7	1.33 (1.30–1.36)	1.37 (1.12–1.67)
Hospital admission						
Misclassified approach						
No monitoring within 1 year	37205	1567	35574	44.0	1.0	1.0
Monitoring within 1 year	36891	932	36193	25.8	0.58 (0.57–0.59)	0.54 (0.49–0.59)
Correctly classified approach						
No monitoring within 1 year	74096	1567	47310	33.1	1.0	1.0
No monitoring within 1 year	37205	1567	35574			
From study entry until monitoring	36891	0	11735			
Monitoring within 1 year	36891	932	24458	38.1	1.15 (1.13–1.17)	1.27 (1.16–1.39)

	Number of patients	Number of events	Person-years	Rate (per 1000 per year)	Crude rate ratio (95% CI)	Adjusted Hazard Ratio†
Drug discontinuation						
Misclassified approach						
No monitoring within 1 year	44871	27875	24704	112.8	1.0	1.0
Monitoring within 1 year	29225	11144	22646	49.2	0.44 (0.43–0.44)	0.46 (0.45–0.47)
Correctly classified approach						
No monitoring within 1 year	74096	27875	32236	86.5	1.0	1.0
No monitoring within 1 year	44871	27875	24704			
From study entry until monitoring	29225	0	7533			
Monitoring within 1 year	29225	11144	15113	73.7	0.85 (0.84–0.86)	1.05 (1.03–1.08)

† Obtained from a Cox proportional hazards time-to-event analysis

5.8.2 Dealing with immortal time bias

I have demonstrated that the person time that was incorrectly allocated to the monitoring group represented over one third of the follow-up time for patients with monitoring. This caused a falsely low rate of events in the patients with monitoring and created an artificial protective effect of monitoring for all three adverse patient outcomes. This tendency for immortal time bias to create an artificial benefit for an exposure has been demonstrated by other authors (Suissa, 2007; Lévesque *et al.*, 2010).

Treating the exposure of interest as time-dependent or time-varying is one method that can be used to control for immortal time bias, and is the method I have used in this analysis. Several additional techniques for dealing with immortal time have been suggested including: (1) using a time matched, nested case-control analysis; (2) prescription time-distribution matching (where the start of follow-up time for treated patients is moved to the end of the immortal period and the follow-up time for untreated patients is assigned according to the distribution of the treated patients' time to start of treatment); and (3) classifying the exposure as a binary value (where use takes the value of 1 for those who started treatment within 90 days of cohort entry and 0 for those who did not). Exposed and unexposed patients are then followed until the event. This therefore follows only 90-day survivors from the same point in time (Suissa, 2005; Zhou *et al.*, 2005; Lévesque *et al.*, 2010).

Immortal time bias is not restricted to studies of drug therapies, but has also been shown to artificially increase the benefit of other interventions. Given the widespread

nature of this bias in published studies, there is a need for increased awareness and improved vigilance in order to control and restrict its effect.

5.9 Discussion

In a large cohort of patients with newly diagnosed hypertension, most patients began treatment with either thiazide diuretics or beta-blockers, consistent with contemporaneous guidelines (Ramsay *et al.*, 1999) and as previously demonstrated in the UK (Walley *et al.*, 2003). In addition, an increase in the proportion of patients treated with thiazide diuretics was observed in 2003, which is similar to the trend seen in North America. Stafford and colleagues (2006) demonstrated an increase in the prescription of thiazide diuretics following the publication of the ALLHAT trial in December 2002, which demonstrated clinical equivalence of thiazide diuretics to ACE inhibitors and calcium channel blockers (The ALLHAT Officers and Coordinators, 2002).

Only 31094 patients (42%) had any baseline biochemical test in the six months prior to beginning antihypertensive treatment, despite guideline recommendations. The absence of baseline testing is inconsistent with standard guidelines. Even older hypertension guidelines state that the measurement of serum electrolyte and urea or creatinine concentrations is ‘essential’ (Sever *et al.*, 1993). These and other simple baseline investigations allow the detection of some causes of secondary hypertension, associated cardiovascular risk factors, evidence of target organ damage, and co-morbid diseases, all of which can influence treatment decisions. Absence of baseline testing also makes assessment of changes in renal function and electrolyte concentrations after the initiation of treatment more difficult.

Half of the cohort had follow-up laboratory monitoring in the year after the initiation of drug therapy; monitoring occurred within six months of starting treatment in a minority (36%) of patients. These rates compare with similar poor rates of follow-up monitoring seen in other studies of antihypertensive drugs used in primary care (Hurley *et al.*, 2005; Clayton *et al.*, 2006b; McAlister *et al.*, 2007; Raebel *et al.*, 2007b) and in hospitalized patients (Uijtendaal *et al.*, 2011). Low levels of monitoring have also been demonstrated in other drugs such as statins (Abookire *et al.*, 2001; Tragni *et al.*, 2007), thiazolidinediones (Graham *et al.*, 2001), and allopurinol (Raebel *et al.*, 2006). A GP practice consultation was recorded within 24 days of treatment for half of the cohort, so the lack of follow-up monitoring was most likely not due to non-attendance but because only a small proportion of patients was monitored.

Overall, less than a quarter of patients had both baseline testing and subsequent follow-up monitoring, depriving their doctors of the opportunity to assess intra-individual changes within the reference range. Also, for those patients who had baseline monitoring alone, this deprives the doctors of the assurance that treatment had caused no biochemical ADR. Guidelines have specifically recommended that all hypertensive patients given an ACE inhibitor should have their renal function measured before and after one week of treatment (Smellie *et al.*, 2007). This is because a rise in previously normal serum creatinine concentration of ≥30% after starting treatment with an ACE inhibitor may suggest renovascular disease. Although more patients treated with an ACE inhibitor had both baseline testing and follow-up monitoring of their serum creatinine concentrations than patients treated with other drug classes, over 60% of those patients did not have their renal function (either creatinine or urea concentration) measured before and after the start of treatment. In

general, only forty percent of patients who did have at least one follow-up monitoring test had a second follow-up test within one year. This lack of additional monitoring limits the ability of GPs to judge sequential changes in serum concentrations.

The proportion of patients monitored after starting treatment increased over the four-year period that I examined. This improvement was similar to the pattern demonstrated in the proportion of patients with a record for weight and smoking in the GPRD, and may reflect a general improvement in standards in general practice (Campbell *et al.*, 2005), or in the way guidelines are implemented. This improvement may have also been as a result of the trend towards the use of incentive schemes targetted at GPs. The Quality and Outcomes Framework (QoF) was introduced as part of the General Medical Services Contract, as a scheme to reward GPs in the UK for how well they care for patients (NHS Employers, 2011). Although QoF was not introduced until after the time period examined in this study—April 2004—the improvement seen in the level of monitoring may have been a reflection of the general trend towards the use of these types of incentive schemes.

Older patients and those with diabetes mellitus were more likely to have either baseline testing or follow-up monitoring, which is consistent with the results of other smaller studies (Hurley *et al.*, 2005; Raebel *et al.*, 2007b). Although patients treated with diuretic drugs and those acting on the angiotensin system were more likely to be monitored than those treated with drugs such as calcium antagonists that have fewer biochemical adverse effects, many patients treated with drugs with a potential to cause biochemical harm went unmonitored.

Patient outcomes

Biochemical ADRs within one year of the start of antihypertensive treatment were rare and the majority of the ADRs were identified solely based on a serum concentration below a certain value and not by a clinical code. Certainly, many patients suffering biochemical disturbances will be asymptomatic and therefore detected only on monitoring. If the patient had not been monitored, the ADR would not have been identified, which may have caused further harm to the patient.

Patients who had any monitoring within the one year of their first prescription were slightly but significantly more likely to be admitted to hospital within that time. This might be explained by admission as a consequence of the results of monitoring tests that showed, for example, hyperkalaemia or important renal impairment. It might also be explained by doctors being more ready to arrange biochemical tests in patients with conditions, such as heart failure, that will themselves require admission to hospital. Alternatively, doctors may be ordering laboratory tests when patients present with an illness that will then progress to a hospital admission. This may be an example of more reactive monitoring or monitoring by indication, rather than a more proactive monitoring approach.

A small proportion of patients died within one year of starting antihypertensive treatment. Patients that were monitored demonstrated a small, although statistically significant, increase in the risk of death compared with those who had no monitoring. However, because of the small number of deaths used in the modelling there is greater uncertainty to the results, and therefore it is difficult to make generalizations to other populations.

Discontinuation of drug therapy is a common problem in clinical practice and constitutes a significant challenge to patients and healthcare providers. Over half of the patients discontinued their first antihypertensive treatment within one year. This proportion is higher than in other studies (Hasford *et al.*, 2002; Burke *et al.*, 2006; Elliott *et al.*, 2007), although this may be due to differences in the distribution or sequence of antihypertensive treatment, the patient population, or the more stringent definition of discontinuation. Of those patients who discontinued first antihypertensive treatment, almost 20% discontinued at 28 days. This would suggest that a large proportion of patients discontinued treatment after one course of treatment. My analysis did not attempt to consider discontinuation rate or its definition as the primary focus. The different discontinuation rates do, however, highlight problems in intermittent compliance during presumed chronic therapy. The discontinuation rate may therefore only be a loose proxy for intolerance to treatment. Monitoring was associated with an increased risk of discontinuation, which may be associated with several factors including increased patient contact, lack of drug efficacy or the serum concentration test detecting an adverse effect of the treatment that would warrant discontinuation.

When I explored the relationship between monitoring and the three outcomes described above, monitoring was treated as a time-dependent covariate in order to control for immortal time bias. The extent of this type of bias is directly dependent on the amount of immortal time that is misclassified (Suissa, 2007). A patient with biochemical monitoring was defined as any patient with a biochemical measurement in the 365 days following the start of antihypertensive treatment. Patients that experienced an outcome shortly after the start of treatment were less likely to be

classified as monitored because the opportunity for having a biochemical test was lower. Likewise, a patient that had biochemical monitoring later (e.g. on the 100th day after starting antihypertensive treatment) would have been free from any outcomes during the first 100 days (thus immortal time). Therefore a majority of events that occurred soon after the start of antihypertensive treatment would have been classified as not monitored as there would have been fewer opportunities to receive biochemical monitoring (Suissa, 2003). This results in a higher rate of the outcomes in the patients that did not have monitoring and artificially decreased the rate of the outcome among patients that had biochemical monitoring, which was seen when I did not model monitoring as a time-dependent covariate.

5.10 Strengths and weaknesses of this analysis

5.10.1 Strengths

This work has examined the nature and frequency of biochemical monitoring using a very large and well-validated database. Most previous studies of monitoring have been carried out using data from American Health Maintenance Organizations, where monitoring is often protocol-driven, and therefore the results of such studies are not as relevant to the UK. The analysis was carried out on a significantly larger population than any of the three previous studies examining monitoring in primary care in the UK (Rhodes, 1992; Kalra *et al.*, 1999; Clayton *et al.*, 2006b). My analysis also looked at both baseline testing and follow-up monitoring and investigated the changes in serum sodium concentrations before and after treatment, which has not been previously carried out in other studies. I undertook modelling to identify patient factors associated with monitoring which importantly adjusted for immortal time bias;

a bias that I have clearly shown can create the illusion of a real benefit in an intervention like biochemical monitoring. Finally, this was the first study that, to the best of my knowledge, examined monitoring not only as a binary outcome, but also explored the frequency and regularity of monitoring.

5.10.2 Weaknesses

There were some limitations and weaknesses to my study. It was impossible to determine the exact context in which monitoring occurred and therefore I cannot be sure that the laboratory tests were undertaken specifically because the patient was being treated with an antihypertensive drug. I was not able to determine from the database whether the GP had undertaken planned or reactive monitoring, in that I could not tease out whether the record of a serum potassium concentration was an indication of a planned monitoring test because the patient was newly treated with an ACE inhibitor, or because the patient presented with an illness that the GP felt required some laboratory tests. Furthermore, I could not determine whether the GP had intended to monitor only potassium, only sodium or both serum electrolyte concentrations, as these tests can generally only be ordered at the same time. Finally, it was impossible to differentiate between GPs failing to undertake biochemical monitoring and patients failing to attend for monitoring.

The GP practice records may not have captured monitoring that may have occured in secondary care. However, most patients with hypertension, especially those with ‘simple’ hypertension requiring therapy with a single agent, are treated in the community. The GP who initiated treatment would usually be responsible for monitoring.

Inevitability, decisions of treatment and monitoring differ amongst GPs. The GPRD does not provide specific information on GPs (e.g. gender, age) or the GP practice itself (e.g. whether it was a group or a single-handed practice). Therefore I was unfortunately unable to make adjustments for potential differences between GPs in their decision and ability to carry out biochemical monitoring due to the lack of available data on GP characteristics.

The majority of patients were included in the cohort based on a code for hypertension and coding for hypertension may not have been consistent across the different practices. Although I used the the threshold for initiation of antihypertensive treatment that was in place at the time, I may also have missed patients treated with antihypertensive agents at blood pressures lower than we defined.

I relied on records of issued prescriptions, and therefore I do not know that the patient collected, or most importantly took, the medicine prescribed (Vrijens *et al.*, 2008). I might therefore have over-estimated the number of patients taking medicines, and so over-estimated the 'non-monitoring' rate.

The proportion of patients who had at least one monitoring test increased significantly during the survey period. This is most likely due to an increase in the number of practices becoming linked electronically to laboratories where tests results are incorporated automatically into the GP practice's electronic records. Without this direct link to the laboratory, it would have been up to the practice to record paper-based laboratory results in their electronic records, which may have caused an under-reporting of monitoring. In addition, the data were collected during a time of change

in United Kingdom general practice that resulted in improvements in the management of hypertension as judged by the increased recording of blood pressure measurements (Ashworth *et al.*, 2008). Although I cannot be sure the data reflect current practice in biochemical monitoring, the differences in monitoring between drug classes and amongst patients are likely to persist.

I chose not to include renal function at baseline as a potential predictor of subsequent monitoring in the models. Although a GP may base their decisions to monitor serum concentrations on a patient's renal function, the inclusion of this variable in the statistical models would have required a measurement of serum creatinine or urea concentration and biased the results by restricting the analyses to patients who already had a laboratory test and might be more likely to be monitored.

Finally, I did not consider potential bias associated with missing data in the study cohort. A significant number of patients were excluded from multivariable analyses examining the patient factors associated with biochemical monitoring due to missing data such as body mass index, which may have led to the introduction of bias into the results. The issue of missing data will be addressed in Chapter 7.

5.11 Conclusions

Just over half over the study cohort had any biochemical monitoring within one year of antihypertensive therapy. Monitoring tended to be sparse and did not tend to follow any regular patterns of testing. The elderly, those with diabetes, and those on ACE inhibitors were more likely to be monitored. Few patients had both baseline tests and subsequent monitoring, but almost one in four of those developed an abnormal test

within one year of starting treatment. The majority of abnormal tests were recorded on the first monitoring test. Despite this, those who were untested were less likely to be admitted to hospital, discontinue treatment or die. The laboratory testing identified a small proportion of patients with evidence of a biochemical ADR. Biochemical laboratory monitoring was therefore able to identify patients who developed an ADR during antihypertensive treatment, as had been supposed. The results suggest, however, that the benefits of monitoring as practiced are slight. It is unclear whether more frequent or more assiduous monitoring in general practice would reduce serious harm from ADRs. There is a need for rational schemes for monitoring ADRs, based on good evidence.

Chapter 6

A SUB-GROUP ANALYSIS OF BIOCHEMICAL ELECTROLYTE MONITORING IN THIAZIDE DIURETIC-TREATED PATIENTS

This chapter presents results from thiazide diuretic-treated patients, who represent the largest sub-group of patients from the initial cohort of patients newly diagnosed with hypertension and newly treated with antihypertensive therapy. Using this sub-group of patients, I further evaluate the nature of sodium and potassium monitoring. I also investigate the frequency of thiazide diuretic-induced hyponatraemia, which is an adverse reaction to treatment that is not well characterized in primary care.

6.1 Introduction

Thiazide diuretics are clinically effective and cost-effective drugs for the treatment of essential hypertension, and are recommended as initial therapy for patients in numerous published guidelines (Chobanian *et al.*, 2003; Guidelines Committee, 2003; National Collaborating Centre for Chronic Conditions, 2006). After the publication of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), which recommended the thiazide diuretic class as the first-choice

antihypertensive medication (The ALLHAT Officers and Coordinators, 2002), this class of drugs became more widely used (Xie *et al.*, 2005). As described previously (chapter 1, sections 1.2.2 and 1.2.3), thiazide diuretics can cause several metabolic ADRs such as hyponatraemia and hypokalaemia.

Hyponatraemia is identified by an abnormally low serum sodium concentration, generally defined as less than 135 mmol/l. Diuretics are the most common cause of drug-induced hyponatraemia, the majority being thiazide diuretics or thiazide diuretic-like agents (Liamis *et al.*, 2008a). Thiazide diuretic-induced hyponatraemia is considered to be of rapid onset, with 50–90% of cases occurring within two weeks of thiazide diuretic use (Spital, 1999; Aronson, 2006; Mann, 2008). Several mechanisms for hyponatraemia due to thiazide diuretics have been postulated including, (1) a positive water balance; (2) negative sodium and/or potassium balance; or (3) a shift of sodium from the extracellular to the intracellular space (Spital, 1999). Mild hyponatraemia is usually asymptomatic, but severe hyponatraemia often requires hospitalization and is associated with symptoms including vomiting, nausea, and lethargy (Sharabi *et al.*, 2002).

Hypokalaemia is often defined as a serum potassium concentration less than 3.5 mmol and has been shown to occur in 10–40% of patients treated with thiazide diuretics (Gennari, 1998). The degree of hypokalaemia due to thiazide diuretic therapy has been shown to be dose-dependent (Ben Salem *et al.*, 2009). In two large randomized controlled trials of low-dose thiazide diuretic treatment, the incidence of hypokalaemia was low (Franse *et al.*, 2000; The ALLHAT Officers and Coordinators,

2002). Patients with low serum potassium concentrations are generally asymptomatic, but the condition can be life threatening when severe.

Hypokalaemia and hyponatraemia can occur concurrently. Indeed, several studies have demonstrated that low potassium concentrations are common in patients with hyponatraemia (Ashraf *et al.*, 1981; Sterns, 1987; Dorup *et al.*, 1988). Potassium deficiency may predispose patients to hyponatraemia because serum sodium concentrations are dependent on the relationship between exchangeable sodium and potassium (Fichman *et al.*, 1971). Hypokalaemia may also increase the risk of serious neurological impairment caused by hyponatraemia (Lohr, 1994).

Most research examining these thiazide diuretic-induced hyponatraemia has been carried out in controlled trials of hospitalized patients, who frequently have significant co-morbidities. The nature and frequency of this ADR is not very well characterized in primary care, where the majority of thiazide diuretics are prescribed (Clayton *et al.*, 2006b). Therefore the aims of this analysis were:

1. To use a large electronic database to examine the nature of electrolyte monitoring in patients treated with thiazide diuretics in primary care;
2. To estimate the frequency of hyponatraemia and concurrent hypokalaemia in this sub-group of patients.

6.2 Methods

Thiazide diuretic-treated patients were identified from the cohort of patients who were newly diagnosed with hypertension and newly treated with a single antihypertensive

agent (described in detail in chapter 3). 1143 women who could have been pregnant at the time of the study were excluded because of potential differences in clinical course, treatment, and monitoring of pregnancy-associated hypertension.

The sub-group analysis focused on the patients who were prescribed a thiazide diuretic alone as their first course of treatment. The same baseline patient covariates described in section 3.3.3 were used in this sub-group analysis. Smoking status, the presence of diabetes mellitus, and body mass index (BMI) were used in the analysis if recorded within five years of the index date, and systolic blood pressure and diastolic blood pressure were used if recorded in the year prior to the index date. I also examined prescription records to see whether carbamazepine or selective serotonin re-uptake inhibitors (SSRIs) were used concomitantly with a thiazide diuretic (within 90 days of starting treatment), or at the time of test for serum sodium concentration. These drugs have been shown to cause hyponatraemia when used concomitantly with thiazide diuretics.

6.2.1 Assessment of electrolyte monitoring and biochemical ADRs

As defined previously in section 3.3.4, I defined baseline sodium or potassium testing as a measurement of one or more serum concentrations between six months prior to the index date and two days following it, to allow for any delays in the uploading of laboratory test results to the GP practice (prior to 2003 most laboratory data entry were entered manually). Biochemical monitoring was defined as a measurement of one or more of these serum concentrations between three days and one year after the index date.

In this sub-group analysis, the occurrence of hyponatraemia or hypokalaemia within one year of thiazide diuretic treatment was identified solely through the use of laboratory tests. Hyponatraemia was defined as a serum sodium concentration less than 135 mmol/l and concurrent hypokalaemia was defined as a record of a serum potassium concentration less than 3.5 mmol/l on same day as the sodium concentration .

6.2.2 Statistical methods

Statistical analyses were carried out using Stata® 10.0. Comparisons were made between data using the Chi-squared test for categorical data, the independent t-test (for data were parametric tests could be used), and the Mann-Whitney *U* test (for data where parametric tests were not appropriate) for continuous data. I determined the relationships between patient variables and the probability of hyponatraemia within one year in univariable analyses, and then entered the baseline characteristics that were statistically significant at the $P<0.05$ level, or were biologically plausible into a backwards stepwise multivariable logistic regression model.

6.3 Results

In the cohort of 74 096 patients who were prescribed a single antihypertensive agent between January 2000 and December 2003, thiazide diuretics was the most commonly used drug class. Almost half of the patients (48%) were newly treated with a thiazide diuretic, of whom 34 877 (97.8%) were treated with bendroflumethiazide—low dose bendroflumethiazide (2.5 mg) was used in the majority of patients (34 472; 98.8%). A

small number of patients were prescribed indapamide (688), chlortalidone (43), and other thiazide diuretics (45).

Almost 40% (13318) of thiazide diuretic-treated patients had a baseline sodium test, with the serum sodium concentration ranging from 117 to 157 mmol/l, (median 140). A similar number of patients (13217) had a baseline potassium test, with a serum potassium concentration ranging from 2.5 to 11.8 mmol/l, (median 4.4). Patients with baseline testing for serum sodium and potassium concentration tended to be male, older, a smoker, and were more likely to have diabetes mellitus (Table 6.1). Of those patients with baseline electrolyte testing, 300 (2.2%) and 218 (1.7%) had evidence of hyponatraemia or hypokalaemia prior to initiation of thiazide diuretic treatment.

Almost one third (11566) of thiazide diuretic-treated patients had at least one sodium concentration measurement in the six months and a one half (16463) in one year following their first thiazide diuretic prescription. Similar rates of potassium monitoring were observed. Men, older and thinner patients, smokers, those with diabetes mellitus, and those that had baseline testing were significantly more likely to have follow-up electrolyte monitoring (Table 6.2).

In those patients with at least one sodium test, the number of tests ranged from 1 to 30 (median = 1 test) (Table 6.3). The time to first follow-up test was highly skewed with a modal time of seven days and a median time of 84 days. Patients with baseline sodium testing had a significantly longer length of time until the first follow-up test and more follow-up tests. Similar results were obtained for potassium monitoring.

Table 6.1 – Baseline characteristics of patients with and without baseline sodium or potassium testing

Characteristic	Any baseline Na^+ testing (N=13318)	No baseline Na^+ testing (N=22335)	P value	Any baseline K^+ testing (N=13217)	No baseline K^+ testing (N=22436)	P value
Male, n (%)	5498 (41.3)	8479 (38.0)	<0.0005	5463 (41.3)	8514 (38.0)	<0.0005
Age, mean (SD)	62.7 (12.3)	63.5 (12.9)	<0.0005	62.8 (12.3)	63.5 (12.9)	<0.0005
<50, n (%)	1920 (14.4)	3257 (14.6)	<0.0005†	1904 (14.4)	3273 (14.6)	<0.0005†
50–59, n (%)	3449 (25.9)	5360 (24.0)		3408 (25.8)	5401 (24.1)	
60–69, n (%)	3710 (27.9)	5888 (26.4)		3698 (28.0)	5900 (26.3)	
70–79, n (%)	3126 (23.5)	5418 (24.3)		3104 (23.5)	5440 (24.3)	
80–89, n (%)	1033 (7.8)	2228 (10.0)		1024 (7.8)	2237 (10.0)	
90–100, n (%)	80 (0.60)	184 (0.82)		79 (0.60)	185 (0.82)	
BMI (kg/m^2), mean (SD)	28.2 (5.3)	28.1 (5.4)	0.445	28.2 (5.3)	28.1 (5.4)	0.240
Underweight, n (%)	131 (1.2)	242 (1.3)	0.202†	130 (1.1)	243 (1.3)	0.151†
Normal, n (%)	3131 (27.4)	5235 (28.2)		3099 (27.3)	5267 (28.3)	
Overweight, n (%)	4645 (40.6)	7370 (39.7)		4612 (40.6)	7403 (39.7)	
Obese, n (%)	3524 (30.8)	5719 (30.8)		3513 (30.9)	5730 (30.7)	
Current or ex-smoker, n (%)	6043 (46.6)	9508 (44.4)	<0.0005	6007 (46.7)	9544 (44.4)	<0.0005
Presence of diabetes mellitus, n (%)	462 (3.5)	596 (2.7)	<0.0005	463 (3.5)	595 (2.7)	<0.0005
Blood pressure mmHg (SD)						
Mean systolic blood pressure	171.2 (18.1)	171.7 (19.1)	0.024	171.2 (18.1)	171.7 (19.1)	0.011
Mean diastolic blood pressure	96.6 (10.5)	96.4 (10.5)	0.113	96.5 (10.5)	96.4 (10.5)	0.350

†Chi squared for trend

Table 6.2 – Baseline characteristics of patients with and without follow-up sodium or potassium monitoring within one year

Characteristic	Any follow-up Na^+ monitoring (N=16463)	No follow-up Na^+ monitoring (N=19190)	P value	Any follow-up K^+ monitoring (N=16477)	No follow-up K^+ monitoring (N=19176)	P value
Male, n (%)	6608 (40.1)	7369 (38.4)	<0.0005	6618 (40.2)	7359 (38.4)	<0.0005
Age, mean (SD)	63.9 (12.4)	62.6 (12.9)	<0.0005	63.9 (12.4)	62.6 (12.9)	<0.0005
<50, n (%)	2121 (12.9)	3056 (15.9)	<0.0005†	2131 (12.9)	3046 (15.9)	<0.0005†
50–59, n (%)	3870(23.5)	4939 (25.7)		3862 (23.4)	4947 (25.8)	
60–69, n (%)	4554 (27.7)	5044 (26.3)		4574 (27.8)	5024 (26.2)	
70–79, n (%)	4253 (25.8)	4291 (22.4)		4256 (25.8)	4288 (22.4)	
80–89, n (%)	1547 (9.4)	1714 (8.9)		1536 (9.3)	1725 (9.0)	
90–100, n (%)	118 (0.72)	146 (0.76)		118 (0.72)	146 (0.76)	
BMI (kg/m^2), mean (SD)	28.0 (5.3)	28.2 (5.4)	0.006	28.0 (5.3)	28.2 (5.4)	0.008
Underweight, n (%)	191 (1.4)	182 (1.2)	0.091†	193 (1.4)	180 (1.1)	0.09†
Normal, n (%)	3990 (28.3)	4376 (27.5)		3989 (28.3)	4377 (27.6)	
Overweight, n (%)	5648 (40.1)	6367 (40.1)		5650 (40.0)	6365 (40.1)	
Obese, n (%)	4273 (30.3)	4970 (31.3)		4285 (30.4)	4958 (31.2)	
Current or ex-smoker, n (%)	7413 (46.1)	8138 (44.5)	0.002	7422 (46.1)	8129 (44.5)	0.002
Presence of diabetes mellitus, n (%)	632 (3.8)	426 (2.2)	<0.0005	635 (3.9)	423 (2.2)	<0.0005
Blood pressure mmHg (SD)						
Mean systolic blood pressure	172.6 (18.9)	170.5 (18.4)	<0.0005	172.6 (19.0)	170.5 (18.4)	<0.0005
Mean diastolic blood pressure	96.4 (10.9)	96.6 (10.2)	0.098	96.4 (10.9)	96.6 (10.1)	0.101
Any baseline testing, n (%)	7096 (43.1)	6222 (32.4)	<0.0005	7031 (42.7)	6186 (32.3)	<0.0005

†Chi squared for trend

Table 6.3 – Summary of serum sodium and potassium concentration monitoring within one year of first thiazide diuretic prescription

	All thiazide-treated patients (N=35 653)	Any baseline Na^+ testing (N=13 318)	No baseline Na^+ testing (N=22 335)	P value		All thiazide-treated patients (N=35 653)	Any baseline K^+ testing (N=13 217)	No baseline K^+ testing (N=22 436)	P value
Any sodium monitoring, n (%)					Any potassium monitoring, n (%)				
Within 1 year	16 463 (46.2)	7 096 (53.3)	9 367 (41.9)	<0.0005	Within 1 year	16 477 (46.2)	7 031 (53.2)	9 446 (42.1)	<0.0005
Within 6 months	11 566 (32.4)	4 358 (32.7)	7 208 (32.3)	0.380	Within 6 months	11 541 (32.4)	4 324 (32.7)	7 217 (32.2)	0.29
Within 28 days	4 854 (13.6)	903 (6.8)	3 951 (17.7)	<0.0005	Within 28 days	4 811 (13.5)	892 (6.8)	3 919 (17.5)	<0.0005
Number of tests in patients with at least one test					Number of tests in patients with at least one test				
Mean	1.62	1.69	1.57	<0.0005‡	Mean	1.62	1.69	1.56	<0.0005‡
Median	1	1	1		Median	1	1	1	
Range	1–30	1–30	1–30		Range	1–30	1–30	1–13	
1 test, n (%)	10 193 (61.9)	4 291 (60.5)	5 902 (63.0)	<0.0005†	1 test, n (%)	10 266 (62.3)	4 270 (60.7)	5 996 (63.5)	<0.0005†
2 tests, n (%)	3 937 (23.9)	1 657 (23.4)	2 280 (24.3)		2 tests, n (%)	3 910 (23.7)	1 636 (23.3)	2 274 (24.1)	
3 tests, n (%)	1 462 (8.9)	698 (9.8)	764 (8.2)		3 tests, n (%)	1 443 (8.8)	689 (9.8)	754 (8.0)	
4 tests, n (%)	508 (3.1)	252 (3.6)	256 (2.7)		4 tests, n (%)	507 (3.1)	243 (3.5)	264 (2.8)	
≥5 tests, n (%)	363 (2.2)	198 (2.8)	165 (1.8)		≥5 tests, n (%)	351 (2.1)	193 (2.7)	158 (1.7)	
Time to first monitoring test (days)					Time to first monitoring test (days)				
Mean	123.6	156.0	99.0	<0.0005‡	Mean	124.3	155.8	100.8	<0.0005‡
Median	84	134	44		Median	84	134	47	
Mode	7	28	7		Mode	7	28	7	

† Chi-squared test for trend; ‡ Comparison obtained through a Mann-Whitney *U* test

In the 16463 patients with a follow-up sodium monitoring test, 9.1% of patients (1505) and 8.4% of the tests (2232/26728) had evidence of hyponatraemia. For the majority of patients (1168/1505; 77.6%), the lowest recorded sodium serum concentration was indicative of mild hyponatraemia (130–134 mmol/l) (Table 6.4). Of those 1505 patients with hyponatraemia, 1410 (93.7%) also had a serum potassium test recorded on the same day. Patients with evidence of severe hyponatraemia had a greater likelihood of being hypokalaemic.

Table 6.4 – Lowest serum sodium concentration recorded within one year and concomitant potassium monitoring in patients with follow-up sodium monitoring

Serum sodium concentration range		Serum potassium concentration test on same day, n (%)	Evidence of hypokalaemia, n (%)
Range	**n (%)**		
≥135 mmol/l	14958 (90.9)	14703/14958 (98.3)	618/14703 (4.2)
130–134 mmol/l	1168 (7.1)	1096/1168 (93.4)	62/1096 (5.7)
125–129 mmol/l	290 (1.8)	273/290 (94.1)	31/273 (11.4)
120–124 mmol/l	37 (0.22)	33/37 (89.1)	5/33 (15.2)
<120 mmol/l	10 (0.06)	8/10 (80.0)	4/8 (50.0)

Ten patients had evidence of very severe hyponatraemia (<120 mmol/l) within one year of thiazide diuretic treatment and their characteristics and electrolyte monitoring are presented in Table 6.5.

Table 6.5 – Characteristics and electrolyte monitoring in patients with severe hyponatraemia within one year of thiazide diuretic treatment

Sex	Age (years)	Baseline Na^+ value	Thiazide , dose (in mg)	Severe hyponatraemia			Monitoring before severe hyponatraemia		Monitoring after severe hyponatraemia	
				Time (days)	Na^+ value	K^+ value	Time (days)	Na^+ value	Time (days)	Na^+ value
Male	71	-	BDZ, 2.5	233	113	4.4	118	130	236	124
							176	131	239	130
Female	74	140	BDZ, 2.5	221	114	-	-	-	225	127
									232	131
									239	130
									267	138
									351	132
Male	83	-	BDZ, 2.5	111	114	2.1	72	132	139	137
									310	137
Female	70	117	BDZ, 2.5	8	116	4.5	-	-	19	123
									38	123
Male	66	138	CHL, 50	95	118	3.0	-	-	109	139
									129	141
									207	138
Male	81	137	BDZ, 2.5	48	118	-	-	-	49	132
									128	132
									139	129
									188	130
									247	132
Male	81	-	BDZ, 2.5	105	118	5.2	10	127	124	120
									147	121
Male	60	-	BDZ, 2.5	7	119	2.7	-	-	14	130
Male	86	-	BDZ, 2.5	12	119	3.3	-	-	335	127
Female	80	142	BDZ, 2.5	151	119	4.0	-	-	-	-

BDZ = bendroflumethiazide; CHL = chlortalidone

The time to the first case of hyponatraemia was highly skewed (Figure 6.1), with a mean number of days of 140, a median time of 113 days, and a modal time of one week. The proportion of sodium tests recorded within 28 days of the first thiazide diuretic prescription that indicated hyponatraemia was 6.9%, which was significantly smaller than the 8.7% of tests carried out later than 28 days after first treatment ($P<0.0005$). Similarly, 6.9% of the tests recorded within 14 days indicated hyponatraemia compared with 8.9% of tests recorded after 14 days of treatment ($P<0.0005$).

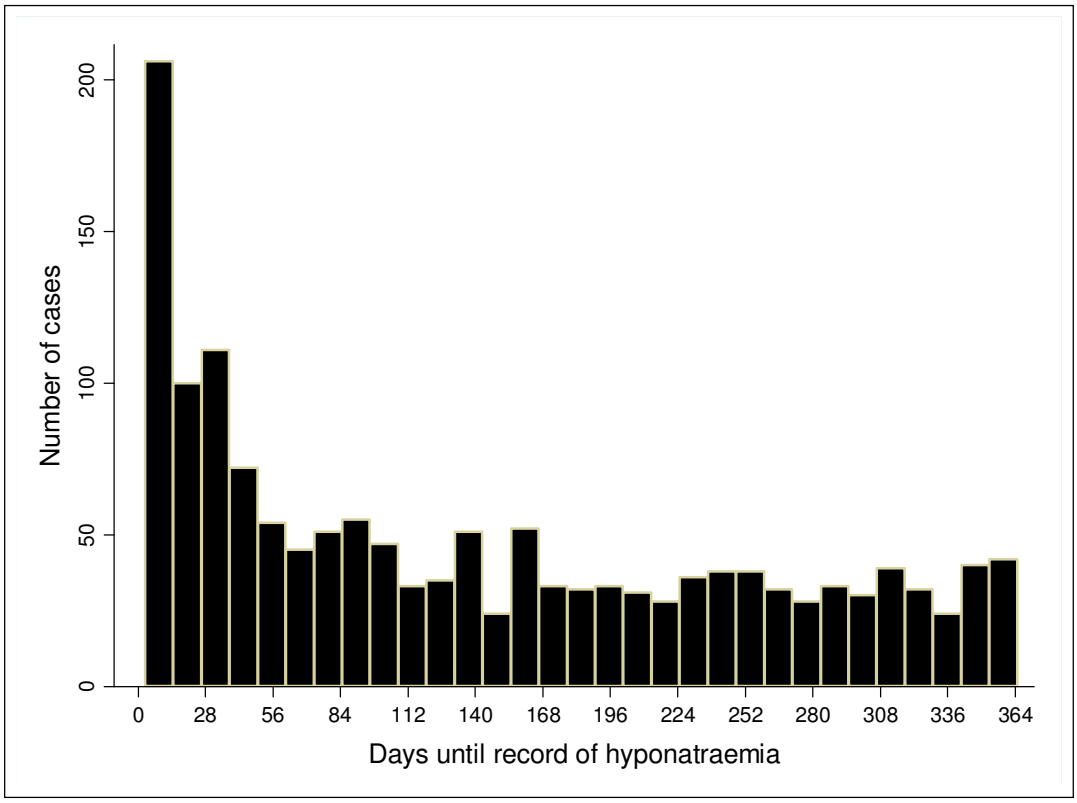


Figure 6.1 – Time until first case of hyponatraemia within one year of thiazide diuretic treatment

The majority of the hyponatraemia cases (1181) were identified on the first serum sodium concentration test. The further sodium monitoring of these patients is described in Table 6.6. A greater proportion of patients with severe hyponatraemia (<125 mmol/l) had a subsequent serum sodium test. The serum sodium concentration normalized in the majority of patients who demonstrated mild hyponatraemia (130–134 mmol/l) on initial testing.

Table 6.6 – Further sodium monitoring and concomitant potassium monitoring in patients with evidence of hyponatraemia on initial testing

Serum sodium concentration on initial testing		No. with a repeat test (%)	Repeat serum sodium concentration		No. with a concomitant K test (%)	Serum potassium concentration	
Range	n (%)		Range	n (%)		Value	n (%)
130–134	955 (80.9)	437 (45.8)	≥135	214 (49.0)	888 (92.9)	≥3.5	835 (94.0)
			130–134	166 (38.0)		3.0–3.4	48 (5.4)
			125–129	50 (11.4)		<3.0	5 (0.56)
			<125	7 (1.6)			
125–129	201 (17.0)	132 (65.7)	≥135	32 (24.2)	185 (92.0)	≥3.5	165 (89.2)
			130–134	68 (51.5)		3.0–3.4	16 (8.6)
			125–129	26 (19.7)		<3.0	4 (2.2)
			<125	6 (4.5)			
<125	25 (2.1)	17 (68.0)	≥135	4 (23.5)	22 (88.0)	≥3.5	16 (64.0)
			130–134	8 (47.1)		3.0–3.4	3 (8.0)
			125–129	3 (17.6)		<3.0	3 (8.0)
			<125	2 (11.7)			

6.3.1 What are the patient factors associated with hyponatraemia?

Women, thinner, older patients, smokers, and those who were prescribed concomitant carbamazepine were significantly more likely to develop hyponatraemia (Table 6.7).

Table 6.7 – Logistic regression model of hyponatraemia within one year of the start of thiazide diuretic treatment

	n/N	Unadjusted OR		P value	Adjusted OR*		P value
Sex							
Male	489/6608	1			1		
Female	1016/9855	1.44	1.29–1.61	<0.0005	1.34	1.17–1.53	<0.0005
Age range							
<50	70/2121	1			1		
50–59	156/3870	1.23	0.92–1.64	0.156	1.11	0.81–1.52	<0.0005†
60–69	355/4554	2.48	1.91–3.22	<0.0005	2.09	1.57–2.79	
70–79	548/4258	4.33	3.36–5.59	<0.0005	3.12	2.34–4.14	

	n/N	Unadjusted OR		P value	Adjusted OR*		P value
80–89	341/1547	8.28	6.34–10.8	<0.0005	5.16	3.80–7.00	
90–100	35/118	12.4	7.79–19.6	<0.0005	6.13	3.26–11.5	
Smoking							
Never	737/8662	1			1		
Ever	700/7413	1.12	1.00–1.25	0.039	1.25	1.10–1.42	0.001
Diabetes mellitus							
No	1416/15831	1			1		
Yes	89/632	1.67	1.32–2.10	<0.0005	1.99	1.55–2.55	<0.0005
BMI							
Underweight	61/191	2.91	2.12–4.00	<0.0005	2.24	1.60–3.14	<0.0005
Normal	554/3990	1			1		
Overweight	406/5648	0.48	0.42–0.55	<0.0005	0.55	0.48–0.63	<0.0005
Obese	162/4273	0.24	0.20–0.29	<0.0005	0.33	0.27–0.40	<0.0005
SSRI treatment at time of test							
No	1453/15889	1			-		
Yes	52/574	0.99	0.74–1.32	0.944	-	-	-
Carbamazepine treatment at time of test							
No	1473/16379	1			1		
Yes	32/84	6.23	4.00–9.70	<0.0005	8.79	5.09–15.2	<0.0005

* Adjusted for sex, age, smoking status, presence of diabetes mellitus, BMI, and concomitant carbamazepine treatment at time of test; † Chi-squared test for trend

6.3.2 Patients with baseline testing and follow-up monitoring

A smaller group of 7096 (19.9%) patients had both a baseline serum sodium test and evidence of follow-up sodium monitoring. A similar number of patients (7031; 19.7%) had both a baseline potassium test and at least one record of follow-up serum potassium concentration within one year. Patients with evidence of hyponatraemia at baseline (191; 2.7%) had a smaller mean percentage change from baseline to the first sodium monitoring test and to the lowest recorded serum sodium concentration (Table 6.8). However, 64.9% (124) of patients with hyponatraemia at baseline had a serum sodium concentration test below 135 mmol/l during follow-up, compared with 8.0% (549) of patients that had a normal baseline serum sodium test (P<0.005).

Table 6.8 – Mean percentage change in serum sodium concentration from baseline to first sodium monitoring test and lowest sodium monitoring test

	All patients (N=7096)	Hyponatraemic at baseline (N=191)	Normal sodium at baseline (N=6905)	P value
% change in serum sodium concentration from baseline to first monitoring test				
Mean	−0.55 (2.11)	+0.27 (3.0)	−0.57 (2.07)	<0.0005
Range	−18.5 to 9.16	−7.52 to 9.16	−18.5 to 7.19	
Median	−0.70	0.75	−0.70	
% change in serum sodium concentration from baseline test to monitoring test with lowest serum concentration				
Mean	−0.95 (2.20)	−0.19 (3.18)	−0.97 (2.16)	<0.0005
Range	−18.5 to 9.16	−7.52 to 9.16	−18.6 to 6.67	
Median	−0.71	0.0	−0.71	

Both the mean percentage change from baseline to the first monitoring test and baseline to the monitoring test with the lowest serum concentrations were significantly correlated with age (P<0.0005; P<0.0005) and gender (P=0.0019; P=0.0283) (Figure 6.2 and Figure 6.3).

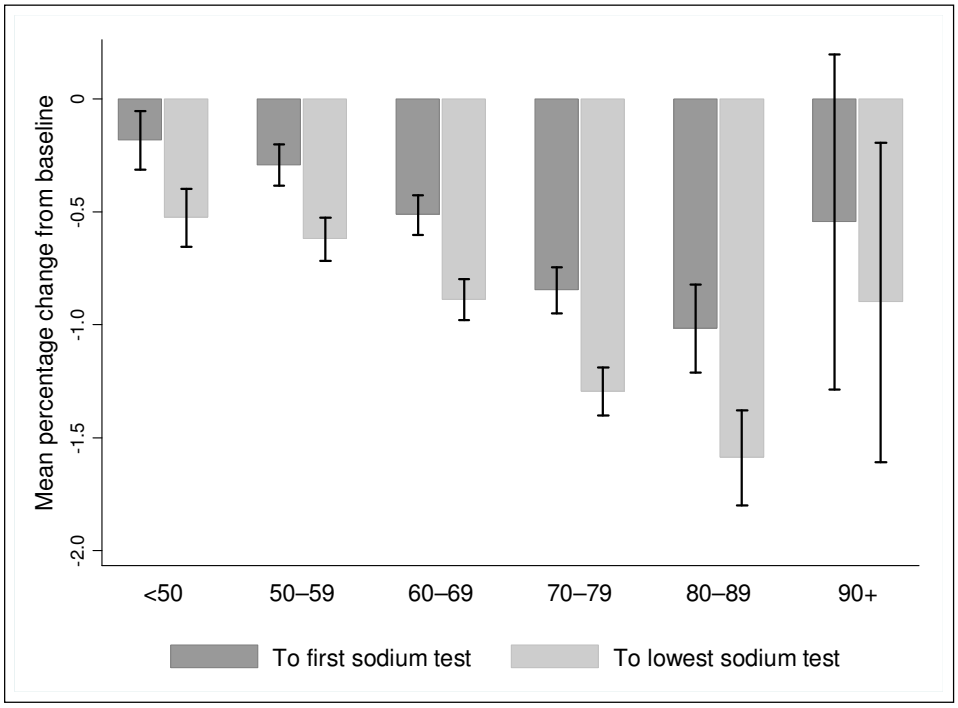


Figure 6.2 – Mean percentage change in serum sodium concentration from baseline to first sodium monitoring test and lowest sodium monitoring test by age group (95% confidence intervals shown) (P<0.0005; P<0.0005)

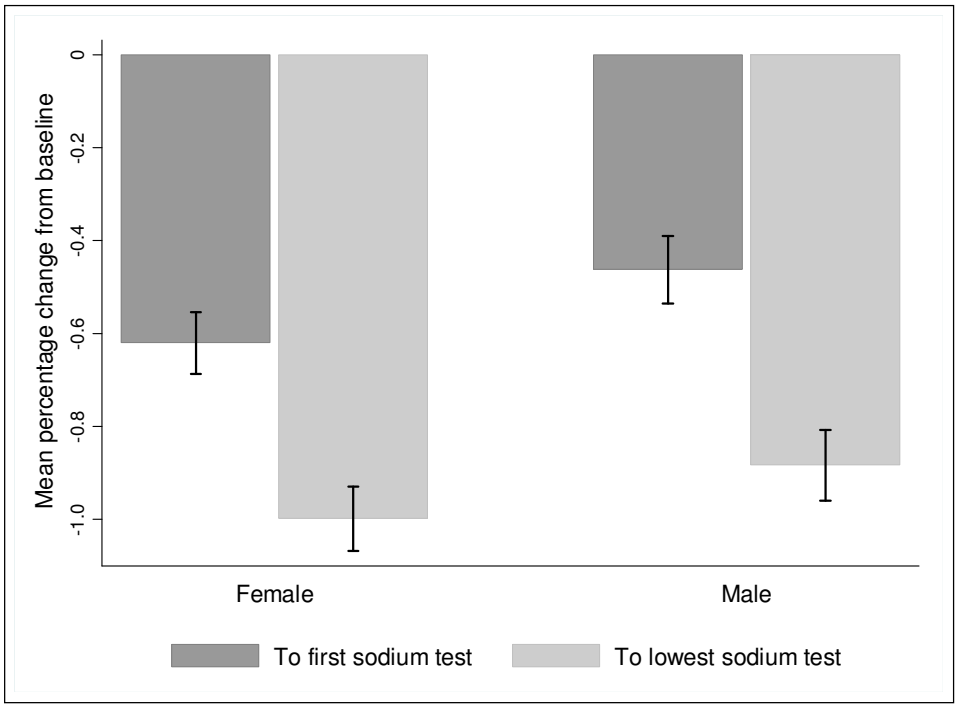


Figure 6.3 – Mean percentage change in serum sodium concentration from baseline to first sodium monitoring test and lowest sodium monitoring test by sex (95% confidence intervals shown) (P=0.0019; P=0.0283)

6.4 Discussion

This analysis of patients newly diagnosed with hypertension confirmed a high level of thiazide diuretic prescribing in UK primary care during the time period studied (48% of patients newly diagnosed with hypertension were treated with a thiazide diuretic). The majority of patients were treated with a low-dose thiazide diuretic—most often bendroflumethiazide 2.5 mg—which was in accordance with published guidelines (Williams *et al.*, 2004).

An examination of thiazide diuretic-treated patients provided specific insight into the nature of electrolyte monitoring and the development of hyponatraemia in a cohort of patients treated with the most common type of antihypertensive therapy. Thiazide diuretics are known to be associated with important biochemical ADRs. Only 37%

had electrolyte testing prior to the initiation of thiazide diuretic treatment and 46% had any electrolyte monitoring during one year of treatment. Patients with baseline electrolyte testing were significantly more likely to have any follow-up monitoring, so too were men, older patients, those with diabetes mellitus, smokers, and those with a lower BMI.

I demonstrated that 9% of patients with serum sodium monitoring had evidence of hyponatraemia during thiazide diuretic treatment. The majority of cases of hyponatraemia following thiazide diuretic treatment were mild. Only 2% of monitored patients demonstrated moderate (125–129 mmol/l) or severe (<125 mmol/l) hyponatraemia. An analysis of subsequent monitoring demonstrated that the majority of cases of mild hyponatraemia normalized on the subsequent serum sodium concentration test. A relationship between hyponatraemia and hypokalaemia was demonstrated as patients with evidence of severe hyponatraemia were significantly more likely to have a serum potassium test indicating hypokalaemia. A small proportion of patients had both baseline testing and follow-up monitoring. Most of these patients (65%) with hyponatraemia at baseline remained hyponatraemic upon subsequent testing in the year following thiazide diuretic prescription, a finding that should perhaps encourage more baseline testing.

The increased risk of hyponatraemia identified in older patients is consistent with the findings of other studies, which have demonstrated that age is an independent risk factor for hyponatraemia (Gross & Palm, 2005; Jiang *et al.*, 2009). This may be related to the decreased ability in older adults to excrete free water as efficiently as younger adults (Clark *et al.*, 1994).

There was an increased risk of thiazide diuretic-induced hyponatraemia in patients who were underweight, although the relationship between low body mass index and hyponatraemia is unclear. Chow and colleagues (2003) suggested that sodium concentration may change more in patients with smaller body size, because they have smaller extra-cellular volume. It is also possible that there is an interaction between smoking, body mass index, and serum sodium concentration: patients with renovascular disease are more likely to be smokers (Mackay *et al.*, 1979; Nicholson *et al.*, 1983), and renal artery stenosis is associated with hyponatraemia (McAreavey *et al.*, 1983).

The increased risk of thiazide diuretic-induced hyponatraemia seen in women is consistent with results from the study of hospitalized patients presented by Sharabi and colleagues (2002), showing that women had a three-fold higher risk of developing thiazide diuretic-induced hyponatraemia. However, this may be due to multiple factors such as a smaller body size, older age, or a lower dietary intake of sodium. The relationship may also be due to an over-representation of women in thiazide diuretic-treated cohorts, instead of an inherent susceptibility to low serum sodium concentrations (Chow *et al.*, 2003; Clayton *et al.*, 2006b).

Patients who were taking carbamazepine were also significantly more likely to develop hyponatraemia. The interaction between thiazide diuretics and carbamazepine to cause hyponatraemia is recognized (Joint Formulary Committee, 2010) and it seems prudent to caution prescribers that monitoring is especially important in patients taking thiazide diuretics with carbamazepine.

Thiazide diuretic-induced hyponatraemia has been shown to occur within 2 to 14 days after the start of treatment (Aronson, 2006). However, the current analysis, like that of Chow and colleagues (2003), failed to demonstrate that hyponatraemia occurs early in thiazide diuretic treatment. Evidence of hyponatraemia was recorded as early as 3 days to as late as 365 days from the start of thiazide diuretic treatment and the proportion of tests with evidence of hyponatraemia was not significantly lower in the first month of treatment.

Limitations

The main limitation of this analysis is the same as those presented in Chapter 5, in that the nature of electrolyte monitoring in this cohort of patients may have been slightly underestimated. This may have occurred due to some monitoring that may have been undertaken in secondary care or that may not have been successfully uploaded into the GP practice's database. There is also the issue of what constitutes a normalized range for serum sodium or potassium concentrations, as the range of concentrations was rather high. GP samples often take hours to reach the laboratory and the 'normal' range is probably different for GPs than for hospital samples.

Another limitation is concerned with the estimation of the frequency and nature of hyponatraemia. The burden of this adverse drug reaction was estimated exclusively using records of laboratory tests. In order to estimate the frequency of hyponatraemia I used the number of patients with serum sodium concentration testing as the denominator and not the number of patients treated with a thiazide diuretic. This may have overestimated the frequency of hyponatraemia seen in thiazide diuretic-treated

patients. It is impossible to determine whether hyponatraemia was missed because the patient did not have a serum sodium concentration test that would have identified the adverse reaction to thiazide diuretic treatment. In addition, because hyponatraemia was identified through laboratory testing it is impossible to disentangle the relationship between testing and the hyponatraemia. In other words, it is impossible to know whether the time to hyponatraemia is an accurate reflection of the time to the event occurring, or simply the time taken for a GP to request a laboratory measurement.

I also presented results that demonstrated significant differences in the patient characteristics between patients with and without baseline sodium testing. The P values showed that the differences were statistically significant and were likely to be real but the magnitude of the absolute differences themselves were very small, and therefore potentially not as clinically important.

Finally, prior research has demonstrated a dose-dependent relationship between thiazide diuretic use and the risk of hyponatraemia. This association could not be explored in this analysis because the majority of patients were treated with the same low dose of bendroflumethiazide.

6.5 Conclusions

The risk of hyponatraemia in thiazide diuretic-treated hypertensive patients is generally low: 9% of the patients with serum sodium concentration monitoring had evidence of hyponatraemia in the first year, of whom only one in five hundred had serum sodium concentrations below 125 mmol/l. Individuals at increased risk of hyponatraemia included women, older patients, patients with a low BMI or diabetes, smokers, and patients taking carbamazepine.

Chapter 7

USING MULTIPLE IMPUTATION IN THE ANALYSIS OF THE GENERAL PRACTICE RESEARCH DATABASE

This chapter describes the problem of missing data in observational research and the biases that can arise when analyses are restricted to patients with complete data. I detail various methods used to deal with missing data and discuss the benefits and limitations of the different methods. I introduce the method of multiple imputation as a tool for dealing with missing data that can reduce the risk of bias. I describe the development of an imputed data set and contrast the range of imputed values with the recorded values. Finally, I compare the results obtained using multiple imputed data with those from a complete case analysis.

7.1 Missing data

Most epidemiological studies contain missing data. Determining whether the missing data will be a problem requires a consideration of the mechanism that caused the missingness. Data can be missing completely at random (MCAR), as in the case where measurement equipment simply malfunctioned in an unpredictable way. When data are MCAR there are no systematic differences between missing values and

observed values. Data can also be missing at random (MAR) where the probability of being missing is not related to the value of the variable that is missing, but may be related to other observed variables (Gorelick, 2006). Finally, data can be missing not at random (MNAR), where the missingness of the data is related to the variable itself and other observed or missing covariates. For example, blood pressure data are MAR if older individuals are more likely to have their blood pressure recorded (and age is included in the analysis), but they are MNAR if individuals with high blood pressure (the variable itself) are more likely to have their blood pressure recorded than other individuals of the same age (White *et al.*, 2010). It is not possible to test whether data are indeed MNAR, as this missing data mechanism depends on unobserved variables. The assumption that data are MAR becomes more likely when more explanatory variables are collected and included in the analysis (White *et al.*, 2010).

7.1.1 Approaches used to determine the mechanism of missing data

One way to determine the missing data mechanism is to create a missing data dummy variable (i.e. data are missing = 0, data are not missing = 1). A significant correlation between the dummy variable and other variables in the data set indicates that missing data are related to other variable(s) and that the data cannot be MCAR. Alternatively, a t-test (for continuous variables) and a Chi-squared test (for categorical variables) can be used to determine if there are any significant differences in the variables between patients who did or did not have missing data. Univariable logistic regression models, where the missing data dummy variable is the outcome variable, can also be used to determine if there is a significant association between patient covariates and data missingness.

7.1.2 Commonly used techniques for handling missing data

The easiest way to deal with missing data is to exclude observations with missing data, which is often called 'complete case analysis'. This is the default for the majority of statistical packages, and is the method I have used in the analyses presented in Chapter 5. This method may result in a large number of cases being discarded, which reduces the study sample size and decreases statistical power. Furthermore, case deletion is also only appropriate when data are MCAR. If the cases that are deleted are systematically different from the rest of the sample, the results of the analysis can be seriously biased (Schafer, 1999).

A method that is commonly used to handle missing data is the indicator method where a new dummy or indicator binary variable for each independent variable is created with 1 indicating a missing value and 0 indicating an observed value. The missing values are recoded as 0 in the original variable. The new indicator variable is then used, together with the original (now recoded) variable, in a multivariable analysis to determine the association between the independent variable and the outcome. If the independent variable with missing data is categorical, an additional category classified as missing is added. This method has greater precision than complete case analysis as no patients are discarded due to missing data. It has been recommended for dealing with incomplete covariate data in randomized control trials because the covariates are not potential confounders due to randomization (White & Thompson, 2005). However, in observational research, this technique will lead to biased associations between the original variables and the outcome, even when the data are missing completely at random (Donders *et al.*, 2006). Furthermore, one cannot adjust for confounding if data from a known confounder are missing.

7.1.3 Simple imputation methods for handling missing data

Simple imputation methods, where a single value for the missing observations is obtained, can also be used to deal with missing data. One such method is mean substitution where each missing data point for a given covariate is replaced by the observed mean value for that patient covariate. Analyses are then carried out in the complete imputed data set. This method has several limitations including the assumption that the data are MAR. Also, the mean substitution will have a mean variance of zero, which may lead to an overestimate in the level of confidence in the imputed data set compared with the observed data set.

Linear regression methods can also be used to fill in or predict the missing data conditional on the basis of other variables. This method allows for the distribution of the variables used in the regression model to be maintained but the variability introduced through the imputation is still ignored as it treats the imputed data as real observed data. This may lead to biased associations and over-precise results.

7.1.4 Multiple imputation for handling missing data

Multiple imputation has been developed and further refined as a method for more appropriately dealing with missing data that uses the distribution of the observed data to estimate a range of plausible values for the missing data. It also adds random components into the model in order to allow for uncertainty from the imputed data. This is repeated several times to created multiple data sets. The estimates are then combined in a way that takes into account the variability of the imputations to obtain the overall estimates, variances, and confidence intervals.

Multiple imputation is a three-stage process. First, *m* values are imputed for each missing data. Missing values are replaced by imputed values that have been sampled from their predictive distribution based on the observed data (Sterne *et al.*, 2009). Imputations are generally undertaken using Markov Chain Monte Carlo or multiple imputation through chained equation techniques. Secondly, the *m* complete data sets are treated as a real complete data set. Each data set will differ because the missing values have been replaced by different imputations. Thirdly, the results of the *m* analyses are combined using Rubin's rules to create a single inference about the parameter of interest that includes a measure of uncertainty from the missing data (Rubin, 1987; Zhou *et al.*, 2001).

Multiple imputation has been shown to perform well when the proportion of the overall missing data is less than 61% (Barzi & Woodward, 2004). The underlying assumption of multiple imputation is that missing data are MAR, and therefore missing values may depend on the observed data but not on the unobserved data. Even if the assumption that the data are MAR does not hold, multiple imputation is less biased than methods such as complete case analysis.

7.1.4.1 Methods for imputing values for the missing data

Markov Chain Monte Carlo technique

Multiple imputations can be created from a multivariate normal model using Markov Chain Monte Carlo (MCMC) techniques. The MCMC method involves simulating draws from a multivariate normal distribution of all of the variables in the imputation model. This method generates predicted values based on the linear regressions and then random draws are made from the simulated error distribution for each regression

equation. The imputed values are created through the addition of the random errors to the predicted values for each individual (Allison, 2009).

Multiple imputations by chained equations

Multiple imputations can also be created through multiple imputation by chained equations (MICE). MICE is a recently developed sequential method where instead of assuming a single multivariate model for all of the data, uses a separate regression model to impute each variable with missing data. MICE cycles through all of the variables, and models each variable conditional on the others (Stuart *et al.*, 2009). Logistic regression is used for incomplete binary variables and linear regression for continuous data.

Therefore unlike MCMC techniques, where values imputed for one variable are never used as predictors to impute other variables, MICE methods use a sequential process so that the values that were imputed in the previous round are then used as predictors for imputing other variables (Allison, 2009). The variable with the least missingness is imputed first, followed by the variable with the second lowest amount of missing data, and so on. Variables with the same amount of missingness are processed in a random order, but the same order is always used (Royston, 2005b). One iteration is complete after all of the variables have been cycled through (Stuart *et al.*, 2009) and the process repeats, imputing missing values until the process reaches convergence (i.e. more iterations will not produce significant changes in the parameter estimates) (Horton & Kleinman, 2007).

MICE techniques may be preferable in certain data sets because MCMC methods assume normality and linearity, and therefore are not well suited for the imputation of categorical variables. MCMC methods are often slow to converge and it is difficult in practice to assess convergence. However, the MCMC approach has stronger statistical underpinnings and there is no theoretical guarantee that the MICE method will converge to the correct distribution for the missing values (Allison, 2009). Recent work by Lee and Carlin (2010) demonstrated that both methods are less biased than complete case analyses and that the results obtained from the two methods are similar.

7.1.4.2 Determining the number of imputations

Some simulation studies have demonstrated that three imputed data sets are sufficient for data where less than 20% are missing (van Buuren *et al.*, 1999). Where rates of missing data are high, more than five to ten imputations tend to have little or no practical benefit (Schafer, 1999). However, more recent research has questioned the claims that more than ten imputed data values are seldom needed. Work by Bodner (2008) has suggested that precision is improved through the use of increasing numbers of imputations. A recent publication has conservatively suggested that the number of imputations should be greater than or equal to the percentage of incomplete cases (i.e. if a data set had 13% incomplete cases, an appropriate number of imputations would be approximately 15) (White *et al.*, 2010).

7.1.4.3 Methods for combining the complete data sets (Rubin's rules)

Once the complete data sets are created, they are then combined using a set of rules in order to obtain the overall estimates, variances, and confidence intervals. This incorporates both within-imputation variability (uncertainty about the results from one

imputed data set) and between-imputation variability (the uncertainty due to the missing data) (White *et al.*, 2010).

The overall multiple imputation point estimate for the parameter of interest is the average of the *m* estimates of the variable *Q* from the imputed data sets:

$$\overline{Q} = \frac{1}{m}\sum_{i=1}^{m}\hat{Q}_i \tag{7.1}$$

The variability of variable *Q* has two components: (1) the estimated within-imputation variance ($\bar{U}$),

$$\overline{U} = \frac{1}{m}\sum_{i=1}^{m}U_i \tag{7.2}$$

and (2) the between-imputation variance (*B*).

$$B = \frac{1}{m-1}\sum_{i=1}^{m}(\hat{Q}_i - \overline{Q})^2 \tag{7.3}$$

The between-imputation variance is the additional variance created by the uncertainty around the missing values. The total variance for the overall multiple imputation estimate is defined as *T*.

$$T = \overline{U} + \left(1 + \frac{1}{m}\right)B \tag{7.4}$$

7.1.4.4 Selection of variables to include in the imputation model

In general, the selection of all available covariates produces multiple imputations with minimal bias and maximal certainty (van Buuren *et al.*, 1999). However, due to computational limitations or problems with multicollinearity, it is often neither feasible nor appropriate to use all variables. A stepwise process for the selection of variables for inclusion into a multiple imputation model using either MCMC or MICE

has been suggested by van Buuren and colleagues (1999). First, one should include all of the patient variables with missing data, the outcome variables and important observed covariates. Failure to include the outcome variable in the imputation of a missing covariate leads to an increase in the risk of bias when determining the association between covariates and the outcome (Moons *et al.*, 2006). Specifically, there is an increased risk of underestimating the covariate-outcome association because there is no covariate-outcome association in the imputed data. Secondly, the variables that are associated with the missingness of the data should be included in the model. Thirdly, variables that are highly correlated with the variables with missing data should be included. Finally, variables that have a very high proportion of missing data should be removed from the imputation model if MCMC methods are being used. Because MICE imputes data variable by variable, one can use different variables that may have been excluded due to a very high proportion of missing data to impute each variable, and therefore MICE can be more advantageous than MCMC methods. The selection of appropriate variables is crucial to providing accurate imputed values.

7.1.4.5 Monitoring convergence

Attempts should be made to determine whether the MICE algorithm has reached convergence or when the chain reaches equilibrium, although no definite method exists. The goal is to have a sufficient number of iterations to stabilize the distribution of the parameters. In general MICE requires fewer iterations than the MCMC methods, which can often require thousands of iterations (van Buuren & Oudshoorn, 1999). Simulation studies have demonstrated that the imputations using a MICE algorithm have stabilized after 10 to 20 iterations, and as few as five (Brand, 1999),

such that the order in which the variables are imputed is no longer an issue (Stuart *et al.*, 2009). One method for assessing convergence of the MICE algorithm is to increase the number of iterations/cycles and examine the data for any noticeable differences. This can be carried out in Stata® by plotting the mean value of each imputed variable against the iteration number (Royston, 2005b). The model has converged when no trend—just random jumps up and down—is apparent in each plot.

7.1.4.6 Imputation diagnostics

After imputed data have been created, one should check to see whether data from the imputations are plausible and whether they differ from the observed data. Differences can arise from the model used to generate the imputations or may indicate that the missingness assumption has been violated, which is a more serious concern (Abayomi *et al.*, 2008). Although there are no agreed tests, statistical and visual diagnostic tests can be used to identify potential problems with the imputed data. A simple graphical method can be to plot the density distribution of observed and imputed values (i.e. only those values actually imputed and not all values in the imputed data sets) (Royston, 2005a; Abayomi *et al.*, 2008). These plots are useful for detecting important differences between the observed and imputed data. Another graphical method is to use bivariate scatter plots, which compare the internal consistency of the imputed and observed observations with respect to a continuous variable. Finally, a significant result from a Kolmogorov–Smirnov test may signal potential differences between observed and imputed values.

7.1.4.7 Assessing the impact of multiple imputation

Finally, one needs to make an assessment of the impact of the multiple imputation. Often comparisons are made between the results obtained using complete case analysis and analyses that have used multiple imputed data.

7.1.4.8 The use of multiple imputation in the literature

The number of publications that have used multiple imputation has increased significantly, although the details of the imputation procedures are often severely lacking (Sterne *et al.*, 2009). Guidelines on the reporting of information on missing data and the implementation of multiple imputation have been suggested by Sterne and colleagues (2009) (Table 7.1).

Table 7.1 – Guidelines for reporting any analysis potentially affected by missing data (from Sterne and colleagues 2009)

- Report the number of missing values for each variable of interest, or the number of cases with complete data for each important component of the analysis
- Clarify whether there are important differences between individuals with complete and incomplete data
- Describe the type of analysis used to account for missing data and the assumptions that were made

For analyses based on multiple imputation

- Provide details of the imputation modelling:
  - Report details of the software used and of key settings for the imputation modelling
  - Report the number of imputed data sets that were created
  - What variables were included in the imputation procedure?
  - How were non-normally distributed and binary/categorical variables dealt with?
  - If statistical interactions were included in the final analyses, were they also included in imputation models?
- If a large fraction of the data is imputed, compare observed and imputed values
- Where possible, provide results from analyses restricted to complete cases, for comparison with results based on multiple imputation.
- Discuss whether the variables included in the imputation model make the MAR assumption plausible
- It is also desirable to investigate the robustness of key inferences to possible departures from the missing at random assumption, by assuming a range of MNAR mechanisms in sensitivity analyses.

7.1.4.9 Software for the application of multiple imputation procedures

Several widely used statistical packages provide methods for the development and analysis of imputed data sets (Harel & Zhou, 2007; Horton & Kleinman, 2007).

- SAS – the PROC MI procedure generates imputed data sets through different methods including MCMC techniques, regression, and propensity score methods. The MIANALYZE procedure combines the results of analyses of imputations.
- Stata® – the *ice* procedure implements multiple imputation by chained equations. The newest version of Stata® 11.0 now has an embedded *mi* procedure that does not implement MI through chained equations. Instead, imputations can be generated through various methods included MCMC techniques.
- R and S-Plus – implement multiple imputation through chained equations

Freely available, stand-alone programmes also exist for undertaking multiple imputation including NORM and MLWin, which use MCMC techniques for the generation of imputed values.

7.1.5 Handling missing data in the General Practice Research Database

The way in which authors have handled the presence of missing data in the GPRD has varied. Often, authors have acknowledged the issue of missing data (Meier *et al.*, 2000; Opatrny *et al.*, 2008), but little work has been done to address potential bias that the significant amount of missing data can introduce to the results. Authors have

sometimes used the indicator method where a missing data indicator was created for each variable with missing data and is then used in regression models (Solaymani-Dodaran *et al.*, 2004; Souverein *et al.*, 2004; Osborn *et al.*, 2007). This method can yield estimates with reduced standard errors relative to a complete case analysis but it can still produce biased estimates (Greenland & Finkle, 1995; Jones, 1996). When data were missing on the length of drug treatment, several authors have also imputed the missing data using the median value from the entire population (Watson *et al.*, 2002; van Staa *et al.*, 2005). This method can also yield biased results.

Very few authors have used multiple imputation with the GPRD, even though the occurrence of missing data, particularly with respect to patient weight and blood pressure, has been widely acknowledged (Gelfand *et al.*, 2005). Five published studies using the GPRD have used multiple imputation to control for potential bias associated with missing data (Delaney *et al.*, 2007; Tannen *et al.*, 2007a; Tannen *et al.*, 2007b; Delaney *et al.*, 2008; Delaney *et al.*, 2009). Delaney and colleagues (2007; 2008; 2009) undertook multiple imputation with MCMC techniques using SAS software. A description of how multiple imputation modelling was undertaken was not provided in either of the studies by Tannen and colleagues (2007a; 2007b).

7.2 Using multiple imputation in the analysis of the GPRD

7.2.1 Introduction

The GPRD, like most other large clinical databases, is prone to missing information. Several methods exist to deal with the missing data, but multiple imputation has been shown to be superior in reducing the risk of biased estimates of effect that can occur

when data are not missing at random. Although the problems caused by missing data and the value of multiple imputation have been previously acknowledged by other authors, the use of this method in the GPRD has been limited.

I have previously examined the relationship between patient factors and biochemical monitoring for ADRs in Chapter 5 using a complete case analysis. However this type of analysis may have biased the estimate of effect. Indeed, one potentially important patient covariate (alcohol use) was excluded entirely from the primary analysis before any regression modelling was undertaken due to the high proportion of missing data.

Therefore the aims of this analysis were:

1. To examine the nature of the missingness of the data in the cohort of patients newly diagnosed with hypertension and newly treated with antihypertensive therapy;
2. To determine if there are important differences between patients with complete and incomplete data;
3. To determine the patient factors associated with biochemical monitoring in a multivariable analysis using multiple imputed data.

7.2.2 Methods

Prior to undertaking multiple imputation, the missing data mechanism was examined. Dummy binary variables were created for each variable with missing data (missing data = 0; non-missing data = 1). The proportion of missing data for each covariate

was also plotted against the year the patient entered the cohort to determine time trends in the measurement of baseline patient covariates.

Multiple imputation was carried out using the Stata® command *ice* (Royston, 2005b), which is an implementation of MICE regression switching algorithm (van Buuren *et al.*, 1999). I did not consider using the MCMC technique to impute the missing data for the practical reason that multiple imputation using chained equations is the method used in Stata® 10.0. A random number seed of 123456 was used to allow for the reproduction of the sets of imputations. Binary variables were imputed using logistic regression and continuous variables using linear regression.

Normality was assumed and was checked prior to undertaking multiple imputation. Continuous variables that did not have a normal distribution were transformed using the *nscore* command (Lunt, 2008), which transformed the data to normality, so that they could be imputed. The *invnscore* command was then used to convert the normally distributed imputed variables back to the distributions of the original variables. The *invnscore* command guarantees that the imputed values cannot lie outside the observed data range. The final step involved rounding the continuous imputed values to the nearest integer.

As a sensitivity analysis, a second method was used to transform continuous variables that did not have a normal distribution. The *lnskew0* command was also used, which created a new variable equal to ln(± skewed variable − K). The command chooses a sign of the variable and a constant K so that the skewness of the new variable is zero.

A back-transformation of the transformed variables was then carried out after multiple imputation, followed by a check to ensure the range of imputed values was valid.

All the patient variables with missing data, the outcome variable and important covariates such as age and sex were selected for inclusion in the imputation model (Table 7.2). Variables that were associated with the missingness of the data, as well as those that were highly correlated with the variables with the missing data, were included in the model. The outcome variable of interest was any biochemical monitoring within one year of antihypertensive and was defined in section 3.3.4. The patient variables were described previously in section 3.3.3.

Table 7.2 – Patient variables used in the development of the imputation model

Variables in the imputation model	Type of variable
Incomplete and outcome variables	
BMI	Continuous
Systolic blood pressure	Continuous
Diastolic blood pressure	Continuous
Smoking status	Binary
Units of alcohol per week	Continuous
Any monitoring within one year	Binary
Covariates	
Age	Continuous
Sex	Binary
Variables related to missingness	
SES quintile	Ordinal
Year of treatment	Ordinal
Drug class	Categorical
8 or more GP visits in year prior	Binary
Any baseline testing	Binary
Prediction variables	
Diabetes status	Binary
PVD	Binary

Ten cycles of regression switching (iterations) were used. For diagnostic purposes, the mean of each imputed variable was plotted against a large number of cycles (e.g. 100). A random pattern would indicate that convergence has occurred. Ten imputations were carried out and the imputed data sets were combined using the Stata® command *mim* (Carlin *et al.*, 2008; Royston *et al.*, 2009). This command can be used to fit regression models in several imputed data sets and can apply Rubin's rules to combine estimates across imputations, as well as calculating the appropriate standard errors for the estimates.

7.2.3 Results

7.2.3.1 Investigating the nature of the missing data

In the cohort of 74 096 patients who were newly diagnosed with hypertension and newly treated with a single antihypertensive medication, smoking status was recorded in 96% of patients, blood pressure in 95% of patients, BMI in 85% of patients, and the number of units of alcohol per week in 59% of patients (Table 7.3). 38 917 (53%) patients were 'complete cases' and had complete data for all the demographic data described below.

Table 7.3 – Amount of missing data in entire cohort

Covariates	Number of records	% missing
Age	74096	0%
Sex	74096	0%
Diabetes status	74096	0%
SES score	74096	0%
Smoking status	71284	4%
Blood pressure	70143	5%
BMI	62691	15%
Units of alcohol per week	43586	41%

Plots of the proportion of missing data by year of first antihypertensive prescription demonstrate that the proportions of missing data for smoking status, blood pressure, BMI, and units of alcohol per week decreased over time (Figure 7.1). The largest absolute decrease was in the proportion of missing blood pressure data, which decreased from 10% to 3% between 2000 and 2003. The decrease in the proportion of missing data was statistically significant for all four variables ($P<0.0005$, Chi-squared test for trend).

The units of alcohol and blood pressure variables did not have a normal distribution and required transformation in order to approximate normality. The number of units of alcohol per week was a semi-continuous variable, in that there was a mixture of zeros and continuously distributed positive values. The zeroes represented actual valid data and were not proxies for negative or missing responses (Schafer & Olsen, 1999).

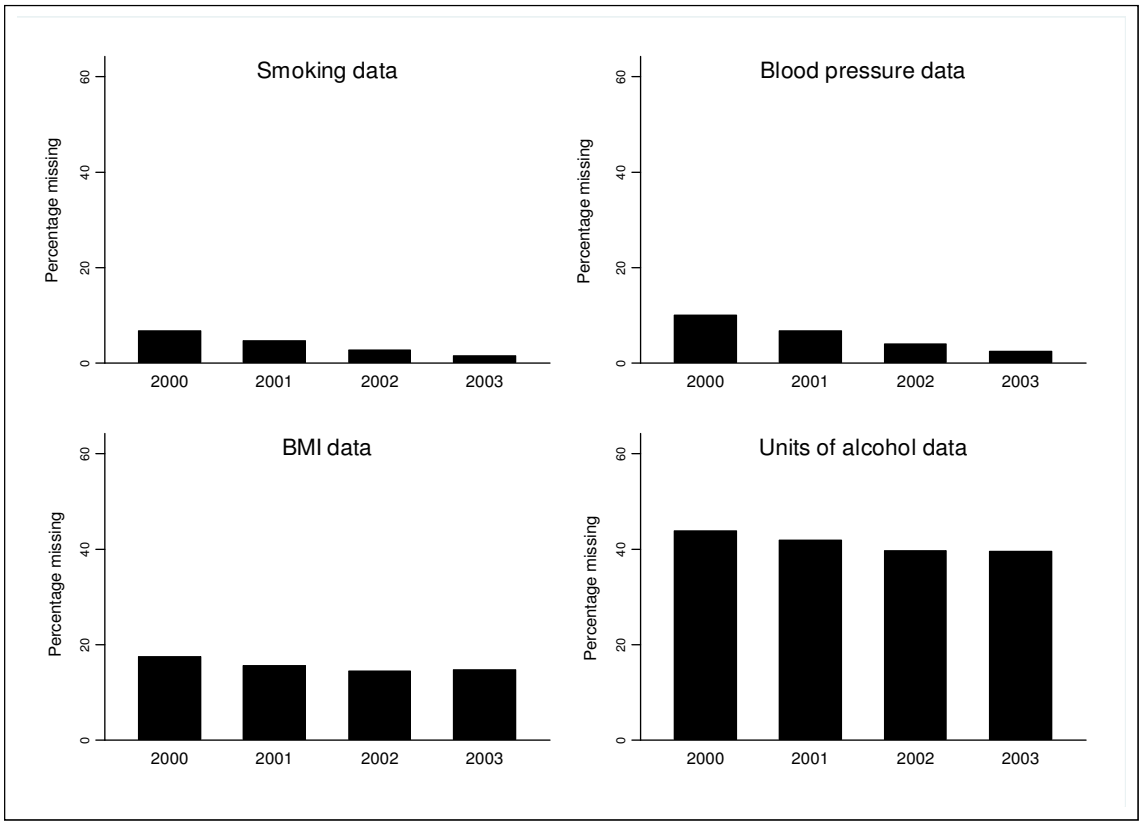


Figure 7.1 – Missingness of baseline patient covariate data by year of first antihypertensive prescription

Univariable logistic regressions demonstrated the presence of relationships between the missingness of the data and several prognostic variables (Table 7.4).

Table 7.4 – Associations between missing data and other patient covariates, and biochemical monitoring

	Missingness of			
	Smoking	Blood pressure	BMI	Units of alcohol per week
Patient covariates				
Age	✓	—	✓	✓
Sex	✓	✓	✓	✓
SES score	✓	—	✓	✓
Diabetes status	✓	✓	✓	—
Year	✓	✓	✓	✓
Outcomes				
Any monitoring within 6 months	✗	✗	✗	✗
Any monitoring within 1 year	✗	✗	✗	✗

✓ = an association ($P<0.05$) was found between missingness of the baseline characteristic and the patient covariate using a univariable logistic regression model; — = no association was found; ✗ = patients with missing data were significantly less likely to have any biochemical laboratory monitoring

Women with missing data for their smoking status were significantly older, were less likely to suffer from diabetes, consumed more units of alcohol per week, had higher systolic and diastolic blood pressures, and had a higher SES score (Table 7.5). A similar pattern was evident in men, although no difference was noted in the SES score.

Significant differences were also observed in patients with missing blood pressure data (Table 7.6). Both men and women with missing blood pressure data were older and less likely to have diabetes. No differences were observed in SES score, BMI, or the recorded number of units of alcohol per week. Male patients with no record of BMI were significantly different in a range of characteristics (Table 7.7). Female patients were only different in their age, diabetes status, SES score, and systolic blood pressure.

Patients with alcohol data were significantly different from those patients who did not have a record of a baseline alcohol use (Table 7.8). Individuals with alcohol data tended to be younger, less likely to have diabetes, more likely to be a smoker and more likely to have a lower SES score.

Table 7.5 – Comparison between patients with and without smoking status recorded

	Male			Female		
	Missing (n=1 240)	**Not missing (n=34 105)**	**P value**	**Missing (n=1 572)**	**Not missing (n=37 179)**	**P value**
Age, mean (SD)	61.2 (14.9)	59.4 (12.3)	<0.0005	67.0 (14.4)	62.3 (12.9)	<0.0005
Diabetes, n (%)	82 (6.6)	3780 (11.1)	<0.0005	53 (3.4)	2 522 (6.8)	<0.0005
SES score, mean (SD)	22.6 (16.8)	22.0 (16.6)	0.21	23.7 (17.1)	22.2 (16.7)	0.0006
Systolic BP, mean (SD)	170.8 (21.1)	169.4 (19.6)	0.03	175.2 (22.2)	172.2 (20.2)	<0.0005
Diastolic BP, mean (SD)	98.5 (11.6)	98.1 (11.6)	0.36	96.9 (11.5)	96.3 (11.0)	0.07
BMI, mean (SD)	29.1 (5.2)	28.5 (4.7)	0.01	28.7 (6.6)	28.4 (5.9)	0.12
Units alcohol/week, mean (SD)	25.0 (25.4)	15.5 (17.0)	<0.0005†	12.3 (21.2)	6.3 (8.8)	0.0003†
Any monitoring within 6 months, n (%)	299 (24.1)	12 879 (37.8)	<0.0005	400 (25.5)	13 368 (36.0)	<0.0005
Any monitoring within 1 year, n (%)	404 (32.6)	17 810 (52.2)	<0.0005	538 (34.2)	18 613 (50.1)	<0.0005

† Mann-Whitney test for non-parametric data

Table 7.6 – Comparison between patients with and without blood pressure recorded

	Male			Female		
	Missing (n=1 942)	**Not missing (n=33 403)**	**P value**	**Missing (n=2 011)**	**Not missing (n=36 740)**	**P value**
Age, mean (SD)	58.7 (12.9)	59.5 (12.4)	0.008	63.2 (14.0)	62.5 (12.9)	0.01
Diabetes, n (%)	151 (7.8)	3 711 (11.1)	<0.0005	99 (4.9)	2 476 (6.7)	0.001
SES score, mean (SD)	22.1 (15.9)	22.0 (16.7)	0.77	22.9 (16.4)	22.2 (16.7)	0.09
Smoking status, n (%)	925 (52.5)	18 106 (56.0)	0.004	685 (37.3)	13 402 (37.9)	0.61
BMI, mean (SD)	28.7 (4.9)	28.5 (4.7)	0.19	28.7 (6.7)	28.4 (5.9)	0.07
Units alcohol/week, mean (SD)	14.7 (16.0)	15.6 (17.1)	0.16†	6.4 (9.6)	6.3 (8.9)	0.51†
Any monitoring within 6 months, n (%)	453 (23.3)	12 725 (38.1)	<0.0005	412 (20.5)	13 356 (36.4)	<0.0005
Any monitoring within 1 year, n (%)	641 (33.0)	17 573 (52.6)	<0.0005	586 (29.1)	18 565 (50.6)	<0.0005

† Mann-Whitney test for non-parametric data

Table 7.7 – Comparison between patients with and without BMI recorded

	Male			Female		
	Missing (n=5 252)	**Not missing (n=30 093)**	**P value**	**Missing (n=6 153)**	**Not missing (n=32 598)**	**P value**
Age, mean (SD)	60.7 (13.5)	59.3 (12.2)	<0.0005	66.6 (13.5)	61.7 (12.7)	<0.0005
Diabetes, n (%)	88 (1.7)	3 774 (12.5)	<0.0005	84 (1.4)	2 491 (7.6)	<0.0005
SES score, mean (SD)	21.0 (17.0)	22.2 (16.5)	<0.0005	21.5 (17.4)	22.4 (16.7)	<0.0005
Smoking status, n (%)	2 373 (53.3)	16 658 (56.2)	<0.0005	1 904 (36.8)	12 183 (38.1)	0.08
Systolic BP, mean (SD)	171.7 (20.7)	169.1 (19.4)	<0.0005	175.5 (21.5)	171.7 (20.0)	<0.0005
Diastolic BP, mean (SD)	98.9 (11.9)	98.0 (11.6)	<0.0005	96.2 (11.6)	96.3 (10.9)	0.39

	Male			Female		
	Missing (n=5 252)	**Not missing (n=30 093)**	**P value**	**Missing (n=6 153)**	**Not missing (n=32 598)**	**P value**
Units alcohol/week, mean (SD)	19.7 (21.1)	15.2 (16.6)	<0.0005 †	8.3 (11.8)	6.2 (8.6)	<0.0005 †
Any monitoring within 6 months, n (%)	1 674 (31.9)	11 504 (38.2)	<0.0005	1 950 (31.7)	11 818 (36.3)	<0.0005
Any monitoring within 1 year, n (%)	2 229 (42.4)	15 985 (53.1)	<0.0005	2 639 (42.9)	16 512 (50.7)	<0.0005

† Mann-Whitney test for non-parametric data

Table 7.8 – Comparison between patients with and without number of units of alcohol per week recorded

	Male			Female		
	Missing (n=11 373)	**Not missing (n=23 972)**	**P value**	**Missing (n=19 137)**	**Not missing (n=19 614)**	**P value**
Age, mean (SD)	60.9 (13.0)	58.8 (12.0)	<0.0005	64.1 (13.3)	61.0 (12.5)	<0.0005
Diabetes, n (%)	1 297 (11.4)	2 565 (10.7)	0.05	1 414 (7.4)	1 161 (5.9)	<0.0005
SES score, mean (SD)	22.0 (16.7)	22.0 (16.6)	0.68	22.9 (17.0)	21.6 (16.4)	<0.0005
Smoking status, n (%)	5 287 (51.6)	13 744 (57.6)	<0.0005	6 007 (34.1)	8 080 (41.3)	<0.0005
Systolic BP, mean (SD)	170.3 (20.3)	169.1 (19.3)	<0.0005	173.3 (20.7)	171.3 (19.9)	<0.0005
Diastolic BP, mean (SD)	97.8 (11.8)	98.3 (11.5)	0.0006	96.0 (11.2)	96.7 (10.8)	<0.0005
BMI, mean (SD)	28.7 (5.1)	28.5 (4.6)	0.0073	28.8 (6.3)	28.1 (5.7)	<0.0005
Any monitoring within 6 months, n (%)	3 963 (34.9)	9 215 (38.4)	<0.0005	6 528 (34.1)	7 240 (36.9)	<0.0005
Any monitoring within 1 year, n (%)	5 458 (48.0)	12 756 (53.2)	<0.0005	9 105 (47.6)	10 046 (51.2)	<0.0005

The significant differences in the characteristics of the patients with and without missing data suggest that the data are not MCAR. However, it is not possible to test whether data are MNAR as this missing data mechanism depends on unobserved variables.

7.2.3.2 Imputation diagnostics

When the mean of each continuous imputed variable was plotted against a very large number of cycles, no tendency for oscillation was observed, which suggested that convergence was achieved and that the ten iterations or cycles were sufficient. The traces of the first 100 cycles of the MICE algorithm are presented in Figure 7.2 to Figure 7.5.

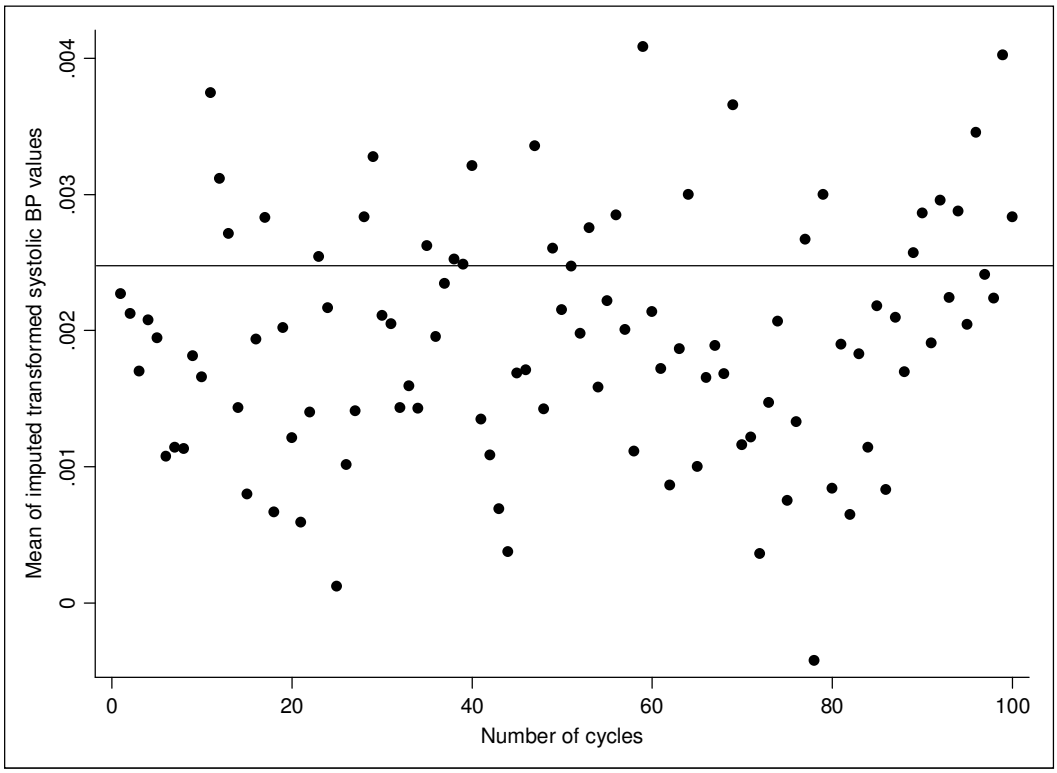


Figure 7.2 – Trace of the first 100 cycles of the MICE algorithm imputing systolic blood pressure (the horizontal line indicates the mean value of the original observations)

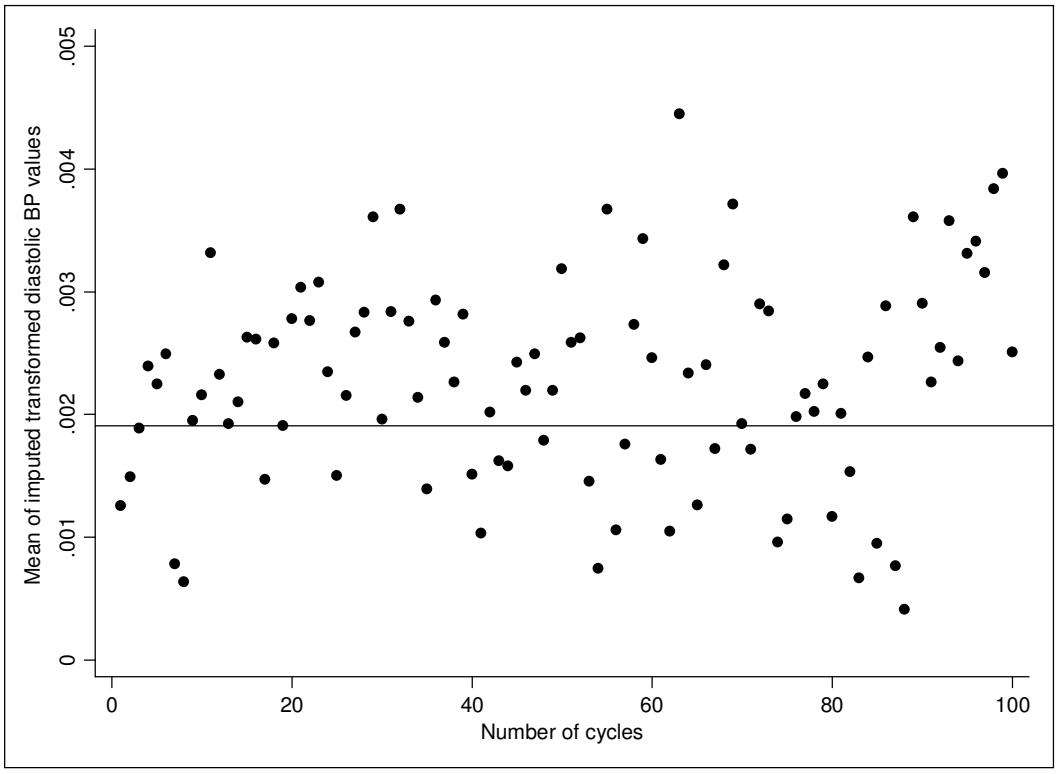


Figure 7.3 – Trace of the first 100 cycles of the MICE algorithm imputing diastolic blood pressure (the horizontal line indicates the mean value of the original observations)

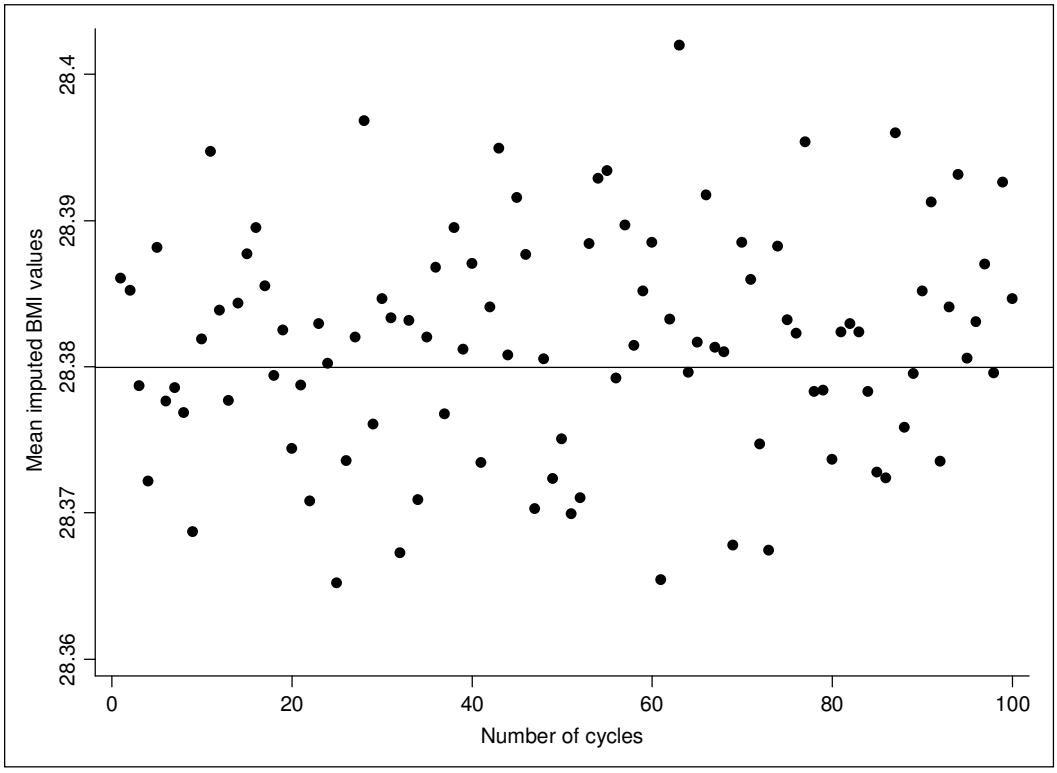


Figure 7.4 – Trace of the first 100 cycles of the MICE algorithm imputing BMI (the horizontal line indicates the mean value of the original observations)

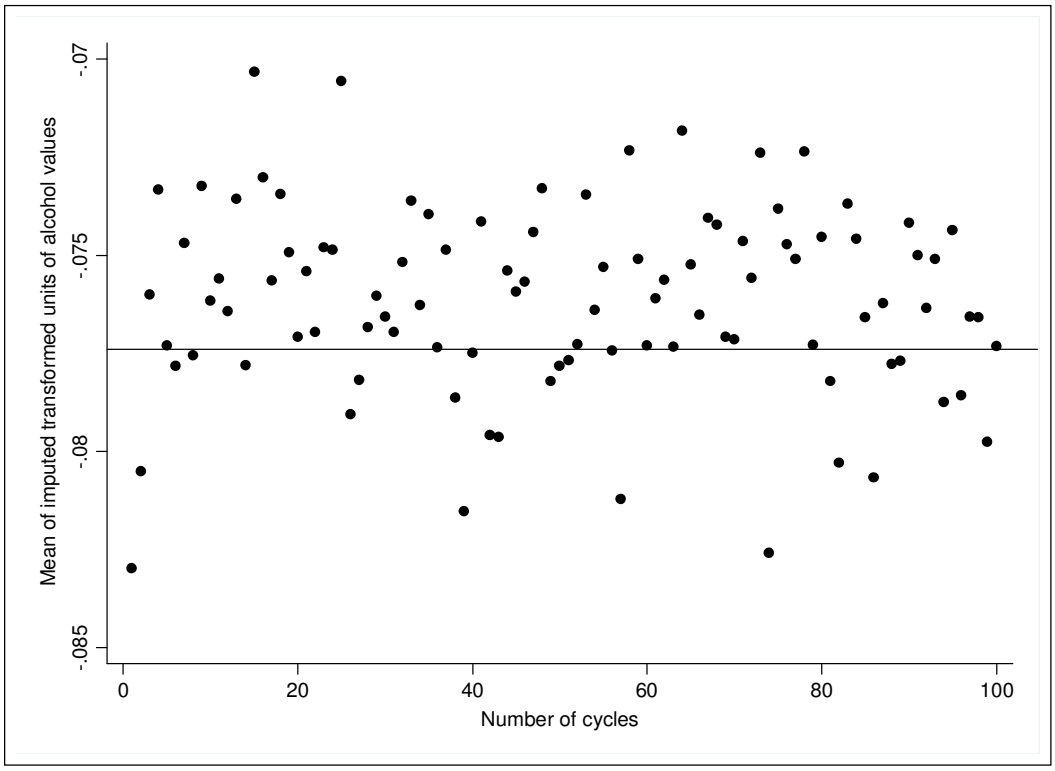


Figure 7.5 – Trace of the first 100 cycles of the MICE algorithm imputing units of alcohol per week (the horizontal line indicates the mean value of the original observations)

7.2.3.3 Comparing observed and imputed values

The summary statistics of the observed and imputed values are shown in Table 7.9. [The term 'imputed values' refers to only those values that were actually imputed and not all values in the imputed data sets, which include both observed and imputed values]. The values of the imputed data were very similar to those from the original data. The summary statistics of the observed data and the complete imputation data sets are presented in Table 7.10.

Table 7.9 – Results from the observed data and the imputed values

	Observed data	Imputed data	Imputed values from the ten imputations									
			1	2	3	4	5	6	7	8	9	10
Smoking status, n (%)	33 118 (46.5)	1299 (46.2)	1294 (46.0)	1295 (46.1)	1322 (47.0)	1338 (47.6)	1228 (45.8)	1306 (46.4)	1296 (46.1)	1267 (45.1)	1338 (47.6)	1250 (46.4)
Systolic BP, mean (SD)	170.9 (20.1)	171.5 (20.6)	171.9 (20.7)	171.6 (20.2)	171.6 (20.8)	171.0 (20.6)	171.2 (20.3)	171.7 (20.8)	171.4 (20.9)	172.0 (20.4)	171.7 (20.5)	171.2 (20.6)
Diastolic BP, mean (SD)	97.2 (11.3)	97.6 (11.5)	97.6 (11.7)	97.6 (11.2)	97.6 (11.3)	97.4 (11.7)	97.5 (11.5)	97.7 (11.8)	97.9 (11.7)	97.6 (11.3)	97.6 (11.4)	97.8 (11.1)
BMI, mean (SD)	28.5 (5.4)	27.9 (5.4)	27.9 (5.4)	27.8 (5.4)	27.8 (5.4)	27.8 (5.5)	27.9 (5.4)	27.8 (5.4)	28.0 (5.4)	28.0 (5.4)	27.8 (5.4)	27.9 (5.4)
Units of alcohol per week, median (IQR)	6 (2–14)	5 (1–14)	5 (1–12)	5 (1–13)	5 (1–14)	5 (1–13)	5 (1–13)	5 (1–12)	5 (1–13)	5 (1–14)	5 (1–14)	5 (1–14)

Table 7.10 – Results from the observed data set and the ten complete imputation data sets

	Observed data	Imputed data	Complete imputation data set									
			1	2	3	4	5	6	7	8	9	10
Smoking status, n (%)	33 118 (46.5)	34417 (46.4)	34412 (46.4)	34413 (46.4)	34440 (46.5)	34456 (46.5)	34406 (46.4)	34424 (46.5)	34414 (46.5)	34385 (46.4)	34456 (46.5)	34368 (46.4)
Systolic BP, mean (SD)	170.9 (20.1)	171.0 (20.1)	171.0 (20.1)	171.0 (20.1)	171.0 (20.1)	170.9 (20.1)	171.0 (20.1)	171.0 (20.1)	171.0 (20.1)	171.0 (20.1)	171.0 (20.1)	171.0 (20.1)
Diastolic BP, mean (SD)	97.2 (11.3)	97.2 (11.3)	97.2 (11.4)	97.2 (11.3)	97.2 (11.3)	97.2 (11.4)	97.2 (11.3)	97.2 (11.4)	97.2 (11.4)	97.2 (11.3)	97.2 (11.3)	97.2 (11.3)
BMI, mean (SD)	28.5 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)	28.4 (5.4)
Units of alcohol per week, median (IQR)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)	6 (2–14)

Figure 7.6 and Figure 7.7 show the density plots and the quintile-quintile plots for the four imputed variables. The distributions of the observed and imputed variables are similar, especially for systolic and diastolic blood pressure, which provides support that the multiple imputation generated values with similar distributions to the observed values.

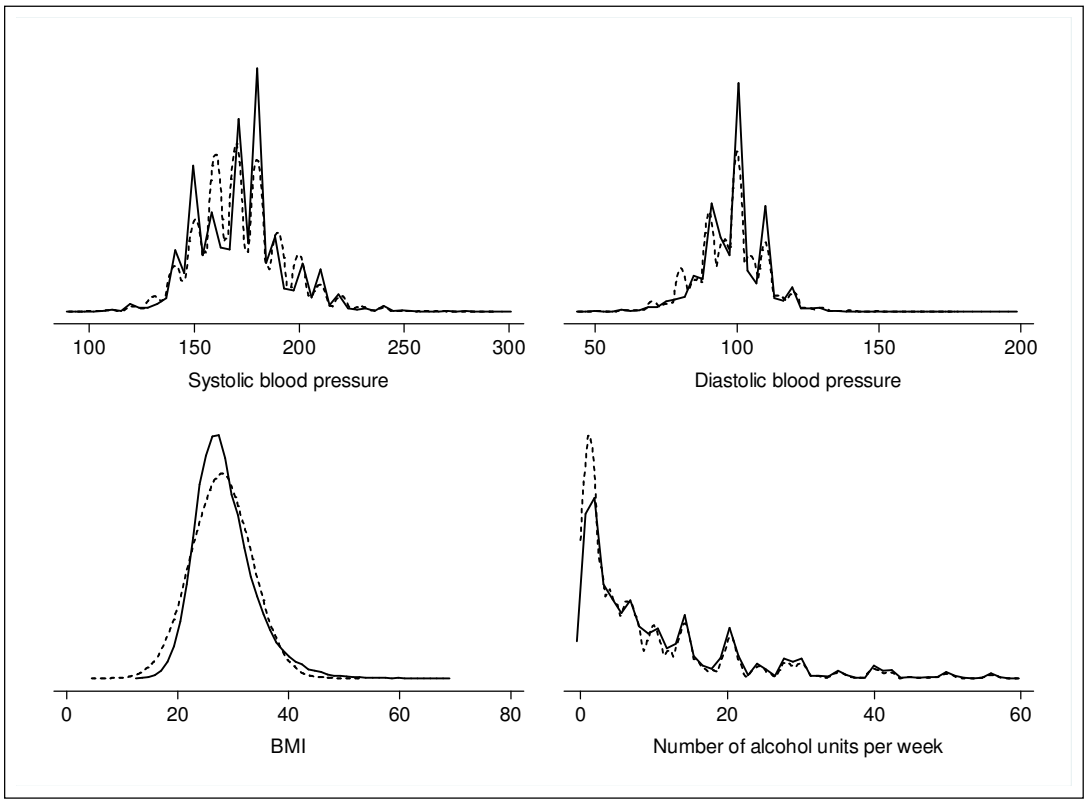


Figure 7.6 – Density plots of observed and imputed variables for four patient variables. For each variable, the solid line shows the density plot of observed values and the dashed line the density plot of the imputed variables.

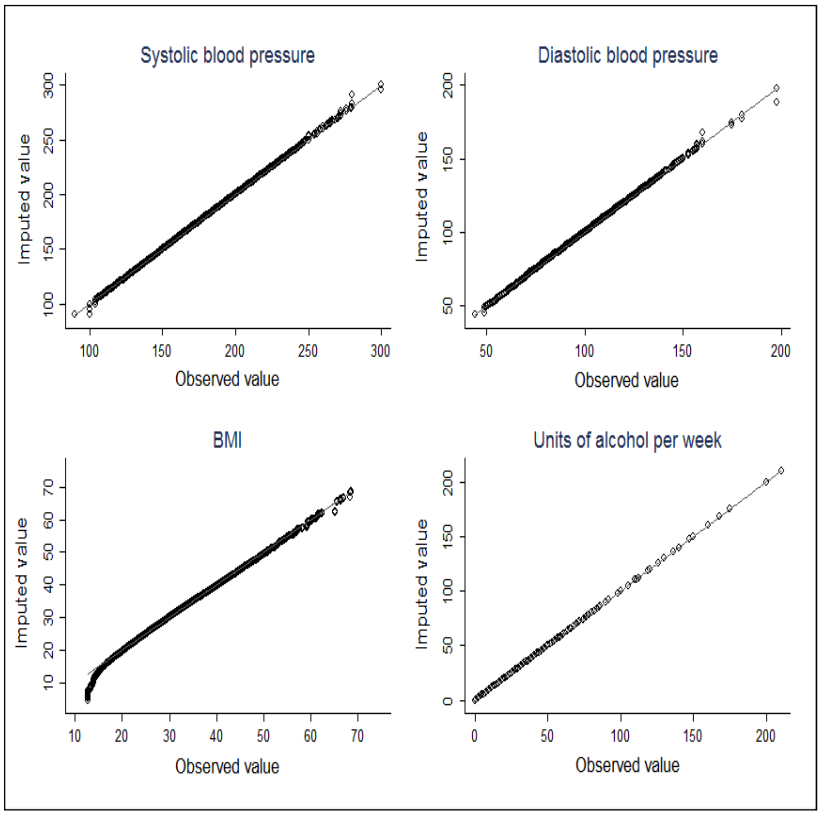


Figure 7.7 – Quantile-quantile plots comparing the distribution of observed values (x-axis) and imputed values (y-axis) for four patient variables

7.2.3.4 What are the patient factors associated with biochemical monitoring using multiple imputed data?

I have previously investigated the patient factors associated with biochemical monitoring, however the analysis was restricted to those patients with complete data and alcohol use was excluded from the outset due to the high proportion of missing data.

Multiple imputation created a complete data set with 74 096 patients; 15 194 more patients than the model using complete case analysis presented in chapter 5 (Table 7.11). Univariable analyses using data from multiple imputation demonstrated the same results as presented in chapter 5: patients with pre-existing diabetes, patients

started on ACE inhibitors at baseline, and those with a first antihypertensive prescription in 2003 were more likely to be monitored. Alcohol use significantly decreased the likelihood of biochemical monitoring. In the multivariable model, monitoring was again significantly associated the presence of diabetes, first-line antihypertensive therapy with ACE inhibitors, and year of first prescription.

The results from the multiple imputed data tended to have narrower confidence intervals compared with the results from the complete case analysis. Differences in the effect size between the results obtained from multiple imputed data and the complete case analysis were seen in some patient factors including older age groups, baseline testing, severe hypertension, and year of initiation of antihypertensive treatment.

As a comparison, a second complete case analysis was also undertaken where alcohol was introduced into the model. This reduced the sample size even more to 38917 patients. The effect size of diabetes mellitus was greater in this data set, while the relationship between hypertension and follow-up monitoring was reduced. Wider confidence intervals were demonstrated in this model, as was expected with the smaller sample size.

Table 7.11 – Logistic regression model of any monitoring in one year following first prescription using multiple imputation methods [results from the previous complete case methods from chapter 5 are also presented as a comparison]

Baseline patient characteristic	Analysis using multiple imputed data (n=74096)						Complete case analysis (n=58902)			Complete case analysis* (n=38917)		
	Unadjusted OR 95% CI		P Value	Adjusted OR 95% CI		P value	Adjusted OR 95% CI		P value	Adjusted OR 95% CI		P value
Sex												
Female	1			1			1			1		
Male	1.08	1.05–1.12	<0.0005	1.03	1.00 – 1.06	0.084	1.02	0.98–1.05	0.323	1.03	0.98–1.07	0.231
Age												
<40	1			1			1			1		
40–49	1.19	1.10–1.28	<0.0005	1.18	1.09–1.28	<0.0005	1.19	1.09–1.30	<0.0005	1.18	1.06–1.31	<0.0005
50–59	1.24	1.16–1.34	<0.0005	1.26	1.17–1.36	<0.0005	1.27	1.17–1.38	<0.0005	1.28	1.15–1.41	<0.0005
60–69	1.48	1.38–1.59	<0.0005	1.47	1.36–1.58	<0.0005	1.52	1.39–1.65	<0.0005	1.55	1.39–1.71	<0.0005
70–79	1.61	1.49–1.73	<0.0005	1.62	1.50–1.75	<0.0005	1.71	1.56–1.87	<0.0005	1.74	1.56–1.94	<0.0005
80–89	1.45	1.33–1.58	<0.0005	1.45	1.32–1.59	<0.0005	1.57	1.41–1.75	<0.0005	1.66	1.45–1.90	<0.0005
90–100	1.25	1.02–1.52	0.033	1.26	1.00–1.57	0.042	1.52	1.10–2.11	0.011	1.46	0.93–2.30	0.103
Smoking status												
Never	1			1			1			1		
Ever	1.08	1.05–1.11	<0.0005	1.05	1.02–1.09	0.001	1.06	1.02–1.10	0.001	1.09	1.04–1.13	<0.0005
Diabetes status												
No	1			1			1			1		
Yes	3.38	3.18–3.58	<0.0005	2.81	2.63–3.00	<0.0005	2.77	2.58–2.97	<0.0005	2.97	2.70–3.25	<0.0005
Alcohol use†												
0–2 (units/week)	1			1			-			1		
3–7	0.94	0.91–0.98	0.004	0.94	0.90–0.98	0.004	-			0.92	0.87–0.97	0.001
8–14	0.91	0.87–0.97	0.002	0.94	0.89–0.99	0.036	-			0.93	0.87–0.99	0.018
>14	0.93	0.88–0.98	0.012	0.94	0.88–1.01	0.077	-			0.96	0.90–1.03	0.271

Baseline patient characteristic	Analysis using multiple imputed data (n=74096)						Complete case analysis (n=58902)			Complete case analysis* (n=38917)		
	Unadjusted OR 95% CI		P Value	Adjusted OR 95% CI		P value	Adjusted OR 95% CI		P value	Adjusted OR 95% CI		P value
SES quintile (0=least deprived)												
0	1			1			1			1		
1	1.09	1.04–1.14	<0.0005	1.08	1.02–1.13	0.003	1.11	1.05–1.17	<0.0005	1.10	1.03–1.17	0.005
2	0.96	0.92–1.00	0.108	0.94	0.90–0.99	0.017	0.95	0.90–1.00	0.063	0.88	0.83–0.94	<0.0005
3	0.90	0.86–0.95	<0.0005	0.91	0.87–0.95	<0.0005	0.94	0.89–0.99	0.016	0.93	0.88–0.99	0.043
4	1.13	1.08–1.18	<0.0005	1.14	1.09–1.20	<0.0005	1.17	1.11–1.24	<0.0005	1.15	1.07–1.22	<0.0005
Hypertension												
Normal or mild	1			1			1			1		
Moderate	0.96	0.93–1.00	0.053	1.10	1.05–1.14	<0.0005	1.07	1.02–1.12	0.007	1.04	0.98–1.11	0.191
Severe	1.13	1.08–1.17	<0.0005	1.32	1.26–1.37	<0.0005	1.22	1.16–1.28	<0.0005	1.19	1.11–1.26	<0.0005
BMI												
Underweight	0.97	0.84–1.12	0.694	0.96	0.83–1.12	0.614	1.02	0.86–1.21	0.805	0.99	0.78–1.26	0.924
Normal	1			1			1			1		
Overweight	1.02	0.98–1.06	0.436	1.00	0.97–1.05	0.807	1.01	0.97–1.05	0.666	1.00	0.95–1.06	0.959
Obese	1.03	0.99–1.07	0.148	1.03	0.99–1.08	0.191	1.05	1.00–1.10	0.029	1.07	1.00–1.13	0.024
Drug Therapy												
Thiazide diuretic	1			1			1			1		
ACE inhibitor	2.00	1.92–2.10	<0.0005	1.57	1.50–1.65	<0.0005	1.62	1.53–1.71	<0.0005	1.58	1.47–1.69	<0.0005
Alpha-blocker	0.93	0.80–1.07	0.314	0.93	0.80–1.08	0.327	0.93	0.79–1.11	0.428	1.00	0.82–1.23	0.965
AT-II receptor antagonist	0.94	0.85–1.04	0.200	0.86	0.78–0.95	0.005	0.89	0.79–0.99	0.047	0.84	0.72–0.97	0.019
Beta-blocker	0.79	0.76–0.82	<0.0005	0.86	0.83–0.89	<0.0005	0.89	0.85–0.92	<0.0005	0.88	0.84–0.93	<0.0005
Ca-channel	0.89	0.84–0.94	<0.0005	0.88	0.83–0.93	<0.0005	0.91	0.85–0.97	0.005	0.87	0.80–0.94	0.001
Loop or K-sparing diuretic	1.05	0.87–1.28	0.598	1.13	0.92–1.38	0.241	1.32	1.02–1.69	0.033	1.15	0.82–1.59	0.420

Baseline patient characteristic	Analysis using multiple imputed data (n=74096)						Complete case analysis (n=58902)			Complete case analysis* (n=38917)		
	Unadjusted OR 95% CI		P Value	Adjusted OR 95% CI		P value	Adjusted OR 95% CI		P value	Adjusted OR 95% CI		P value
Combination	0.70	0.60–0.81	<0.0005	0.73	0.63–0.85	<0.0005	0.83	0.70–0.99	0.033	0.84	0.68–1.04	0.109
Any baseline testing												
No	1			1			1			1		
Yes	1.48	1.44–1.53	<0.0005	1.20	1.17–1.24	<0.0005	1.15	1.11–1.19	<0.0005	1.12	1.07–1.17	<0.0005
Year of first prescription												
2000	1			1			1			1		
2001	1.35	1.29–1.41	<0.0005	1.34	1.28–1.40	<0.0005	1.31	1.25–1.38	<0.0005	1.30	1.22–1.38	<0.0005
2002	1.90	1.82–1.98	<0.0005	1.91	1.82–2.00	<0.0005	1.83	1.74–1.92	<0.0005	1.78	1.68–1.89	<0.0005
2003	2.89	2.77–3.02	<0.0005	2.92	2.79–3.05	<0.0005	2.80	2.66–2.94	<0.0005	2.68	2.52–2.84	<0.0005

† Units per week are presented for females. Alcohol use in males was categorized as: 0–3; 4–14; 15–21; >21 units per week.; * complete case analysis with alcohol used entered into the model

7.2.3.5 Sensitivity analysis of the transformation of continuous variables

I used a user-written Stata® command (*nscore*) to transform the variables that did not approximate a normal distribution. As a sensitivity analysis, I used another Stata® command (*lnskew0*) to transform the continuous variables. Density plots demonstrated that the distributions of the observed and imputed variables generated using the *lnskew0* function were similar (Figure 7.8).

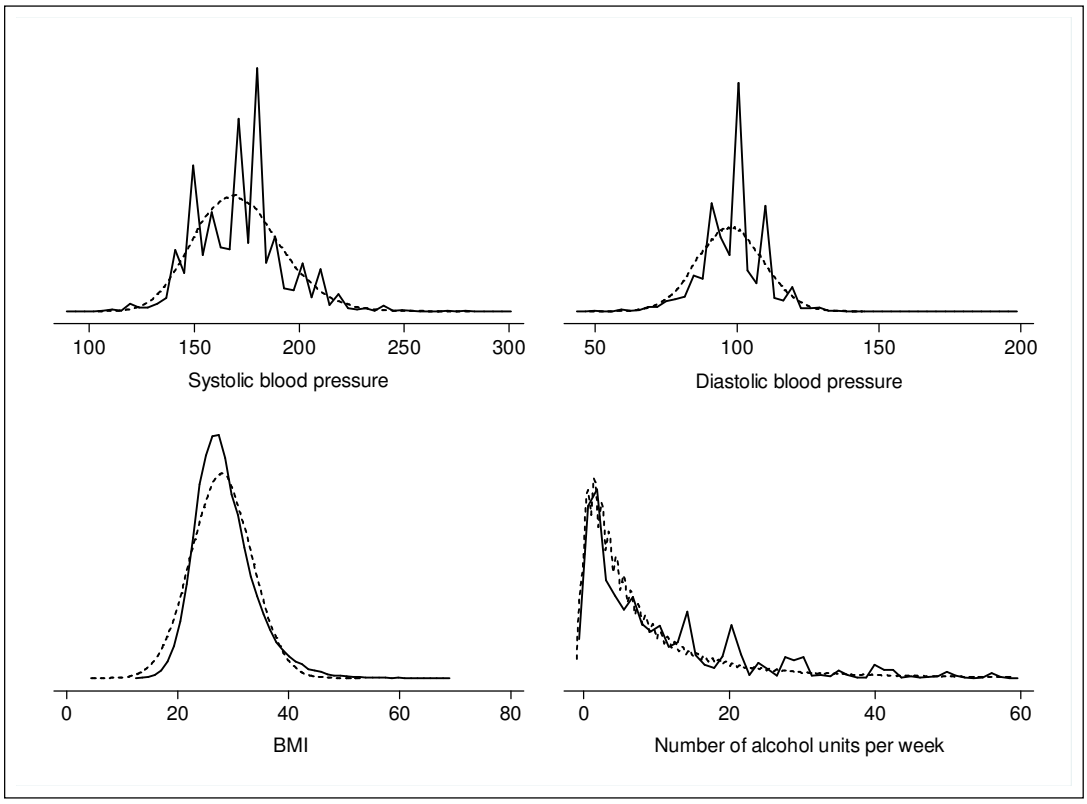


Figure 7.8 – Density plots of observed and imputed variables for four patient variables using the lnskew0 function. For each variable, the solid line shows the density plot of observed values and the dashed line the density plot of the imputed variables.

The imputed values generated by both the *nscore* and *lnskew0* Stata® commands generated similar results. The multiple imputed data generated by the *lnskew0* command also generated similar results in the multivariable analysis investigating the

patient factors associated with monitoring as those data generated by the *nscore* command.

7.3 Discussion

Primary care databases such as the GPRD are recognized as rich data sources, which provide for opportunities to analyse various exposures and populations that would potentially be impossible or unethical in clinical trials (e.g. studies of the elderly or in pregnant women). These databases often provide a good number of important patient covariates such as blood pressure, height, and weight. However, the proportion of missing values for these variables can be significant, which can lead to potential issues in the analysis and the significant potential for bias.

Complete case analysis, where records with missing data are discarded, is an easy method for dealing with the problem of missing data and is the method used in the analyses I presented in chapter 5. This method does, however, exclude variables with large quantities of missing data (e.g. alcohol use), which does not allow for the ability to adjust for potential confounding by these variables. Complete case analysis also reduces the precision of the effect estimates and can introduce significant bias if the assumption that the data are missing completely at random does not hold. Indeed, upon examination of the missingness of the data, I found that there were significant differences in the patients with and without missing data, and therefore demonstrated that the data were not missing completely at random.

Multiple imputation is a powerful technique that is used to deal with missing data, that does not require data to be missing completely at random. The use of this

technique has increased significantly, with several statistical packages providing options to both create and to later analyze the multiple imputed data. Multiple imputation has been shown to be successful in accounting for missing data on weight, height, and blood pressure in primacy care databases (Marston *et al.*, 2010). Although the same authors cautioned that multiple imputation may not be suitable for imputing alcohol use or smoking status as the assumptions used in multiple imputation may be violated.

I used multiple imputation by chained equations to create a data set with an additional 15 194 patients and use all of the information in the incomplete records. This method prevented the loss in power from having to exclude a patient that is missing only one variable. Multiple imputation by chained equations specifies different regression models for each incomplete variable, which is beneficial because it does not assume a specific form for the multivariate distribution. This method does assume that the regression models converge to the joint multivariate distribution. Convergence was achieved, which suggested that the number of iterations used was satisfactory.

I was able to create a range of imputed values with distributions that were similar to the observed values using multiple imputation. The results from the analysis using imputed data were similar to those from the complete case analysis, with some differences in the magnitude of the effect of some patient factors on monitoring. For example, the absolute difference in the association between age and biochemical monitoring in the multiple imputed data set was 12% lower than in the complete case analysis. This may have been caused by differences in how data were recorded in

these older age groups, which may have led to data being missing and the subsequent exclusion of these patients from the complete case analysis.

Multiple imputation is based on the assumption that the data are MAR, and departures from MAR, which tend to occur with large amounts of missing data, can affect the results. This can be reduced through the inclusion of many variables in the imputation model in order to make MAR more plausible. The selection of the variables into the imputation model was based on practical and theoretical considerations in that variables that predicted the missingness of a missing variable or were correlated with a missing variable were selected. The database consisted of hundreds of variables that may have been used in the imputation model. However, the inclusion of all of the variables would not have been computationally feasible and may have led to instability of the model.

7.4 Conclusions

Imputed data will never be a substitute for observed data, and steps should be taken by researchers to minimize the proportion of data that are missing. However, multiple imputation by chained equations was able to generate plausible values for variables with missing data that were similar to the observed values. When I examined the patient factors associated with biochemical monitoring using multiple imputed data, similar results were obtained to those from the previous complete case analysis. This does not provide clear evidence that the complete case analysis is more appropriate. Indeed, the results using the multiple imputed data had narrower confidence intervals and more precision.

Chapter 8

ESTIMATING THE RELATIONSHIP BETWEEN BASELINE BIOCHEMICAL TESTING AND ADVERSE PATIENT OUTCOMES USING PROPENSITY SCORE METHODS

In this chapter I introduce the propensity score as a tool for controlling confounding in observational research. I discuss the increase in the use of propensity score methods and their potential benefits and risks. I describe the development of a propensity score to estimate the probability of baseline biochemical testing, which is then used to examine the relationship between biochemical testing and adverse patient outcomes. Finally, I compare the results obtained using propensity score methods and traditional multivariable models.

8.1 Controlling for confounding in observational research

8.1.1 Comparing randomized controlled trials with observational studies

Randomized controlled trials (RCTs) are viewed as the gold standard for estimating treatment effects. Randomization in RCTs ensures that measured and unmeasured patient covariates are balanced in both treatment arms, which reduces the risk of

confounding. However, RCTs can sometimes be unnecessary, inappropriate, impossible, or inadequate (Black, 1996). For example, a RCT may be unethical or unfeasible when one wishes to study the effects of treatment in certain patient groups such as pregnant women or children. RCTs are also often too small to identify rare adverse events; a limitation that is addressed through the use of post marketing surveillance schemes to identify adverse drug reactions such as the Yellow Card Scheme in the UK (Davis *et al.*, 2007). RCTs also commonly enrol a sample of patients that may not be representative of the population of interest due to restrictive selection criteria. Patients enrolled in a RCT tend to be healthier, younger, and more likely to be male (Hennekens & Buring, 1987). RCTs are also limited by their short length of duration, which can limit the ability to detect delayed adverse reactions to treatment. Finally, a randomized experiment may not be appropriate if randomization itself will reduce the effectiveness of an intervention. This will occur if the participation of the patient was dependent on the patient's own preferences (e.g. health promotion or disease prevention programmes) (Black, 1996).

Because RCTs are often impossible due to ethical, logistical, or practical limitations described above, observational studies are required. Observational studies can provide an assessment of real world practices with regards to the behaviours of both physicians and patients. Observational studies, particularly those that are retrospective in design, allow for the timely and cost-effective assessment of treatments. They also can provide opportunities to analyze long-term follow-up data that would have not been feasible with a RCT.

The main source of potential bias in observational studies is selection bias, where the treatment groups may differ systematically at baseline (Deeks *et al.*, 2003). In a randomized experiment, patients are allocated to treatment groups through chance. For observational research, allocation to treatment will be influenced by a range of factors including a doctor's experience and various patient factors. This tendency can result in treatment groups that are systematically different and impossible to compare. Confounding by indication—where differences between two groups originate from differences in the indication for treatment such as the presence of various risk factors—is of specific concern. The decision to treat a patient with a certain intervention can be influenced by some factor that is associated with the outcome of interest. Confounding by indication can lead to an overestimation or underestimation of the true association between a treatment or exposure and an outcome (Rothman & Greenland, 1998).

Observational studies are also at risk of (1) attrition bias, which is associated with the completeness of follow-up; (2) detection bias if the assessment of the events of interest is not blinded; and (3) performance bias if there are systematic differences in the way that an intervention is allocated, applied, and recorded (Deeks *et al.*, 2003).

8.1.2 Controlling for confounding in observational studies

Various methods have been suggested for controlling for measured confounding in observational studies (Rothman & Greenland, 1998; Greenland & Morgenstern, 2001). One method is *restriction*, where analyses are carried out in a sub-sample of patients that are alike with regards to the confounding variable. For example, if smoking is thought to be a confounder then the analysis can be carried out solely in

non-smokers. This method removes the potential for confounding based on smoking, but can significantly decrease the sample size and limit how one extrapolates the results to other groups.

Matching based on one or more of the suspected confounders can also be carried out to control for confounding. Exposed patients are matched with unexposed patients (in the case of a cohort study) based on the suspected confounding factor (e.g. age or sex). Matching can control for confounding more efficiently with less random error in the statistical analysis (Greenland & Morgenstern, 2001). However, it may become difficult to obtain matches if attempts are made to match on more than one confounding factor.

Standardization is a method of data analysis that combines information over strata using a specified system of weights. Weights from a standard population are used to determine the number of events that would have been expected if the two populations being compared had had identical distributions of the potential confounding factor (Hennekens & Buring, 1987). When rates adjusted for a particular factor are compared, any remaining observed difference between the groups cannot be attributed to confounding by that factor.

Stratification is another data analysis method that can control for measured confounding. The results are stratified by levels of one or more confounding factors and the effect of the exposure is estimated with a given stratum. Because the confounding factor is homogeneous within each stratum, it cannot act as a confounder. The effect is then summarized across all of the strata through a pooled

estimate. The method is limited to one or two potential confounders and can lead to problems with sparse data as few strata may have both treated and untreated patients with all of the confounding factors. Another disadvantage is that continuous variables will have to be classified into strata using potentially arbitrary criteria (Martens, 2007).

Finally, *multivariable regression* modelling examines the relationship between an exposure and an outcome while simultaneously controlling for multiple confounding factors. The exposure is treated as the independent variable of interest and the analysis adjusts for the treatment effect of all of the confounders added into the model. This method adjusts for significantly more confounders than stratification, but there is a limit to the number of potential confounders that can be modelled and still generate stable estimates. A rough rule of thumb of no fewer than ten events or outcomes per variable has been suggested in order to maintain the validity of the model and reduce the occurrence of biased estimates (Peduzzi *et al.*, 1995; Harrell *et al.*, 1996; Peduzzi *et al.*, 1996), although this 'rule of ten' may be too strict (Vittinghoff & McCulloch, 2007). The ease of use of multivariable modelling has increased substantially with improvements in computing systems and statistical software. However, the ability of multivariable modelling to control for confounding is only as good as the model itself. Indeed, Rothman (1986; p. 285) stated that:

> 'The first experience with multivariate analysis is apt to leave the impression that a miracle in the technology of data analysis has been revealed; the method permits control for confounding and evaluation of interactions for a host of variables with great statistical efficiency. Even better, a computer does all the arithmetic and neatly prints out the results. The heady experience of commanding a computer to accomplish all these analytic goals and the simply gathering and publishing the sophisticated 'output' with barely a pause for retyping is undeniably alluring... However useful it may be, multivariate analysis is not a panacea...Multivariate analysis provides a way to preserve precision while controlling many variables, by postulating a mathematical

> model that allows the data to be used more efficiently to estimate many effects simultaneously. The extent to which this process represents improved efficiency rather than just bias depends on the adequacy of the assumptions built into the mathematical model.'

8.2 A summary of propensity score methods

Propensity score methods have also been suggested as a way of controlling for measured confounding. The propensity score is a single, unidimensional multivariate score calculated for each patient within a given cohort that estimates the patient's chance of receiving treatment according to their characteristics. The use of a multivariate score to control for measured confounding was first suggested by Miettinen (1976) over thirty years ago. He described an approach where confounding factors are summarizing using a single, unidimensional score. Rosenbaum and Rubin (1983; 1984) built upon Miettinen's work, coining the term 'propensity score', and advocating the use of propensity score methods as a way of controlling for confounding in observational studies. The use of propensity score methods to estimate treatment effects using observational data has increased significantly, with a dramatic increase in the past few years (Weitzen *et al.*, 2004; Stürmer *et al.*, 2005; Stürmer *et al.*, 2006a; Austin, 2008b).

8.2.1 Estimating the propensity score

The propensity score is defined as the conditional probability of an event—usually the treatment—given an individual's observed covariates (D'Agostino, 1998). A propensity score captures how differences in these observed covariates contribute to the probability of a patient receiving a given treatment. The score has a value between zero and one, so each patient has a non-zero probability of receiving treatment. The

goal of the propensity score model is not to predict patient allocation into treatment groups accurately but to create a score that will create a balance on covariates over groups within propensity score strata (Brookhart *et al.*, 2006). For a given propensity score, the distribution of the covariates defining the propensity score is on average the same in the treated and the untreated groups, thus creating a quasi-randomized experiment, at least with respect to these covariates (Stuart & Rubin, 2008).

The propensity score is most often estimated from a logistic regression model where the event (most often drug treatment) is a binary outcome (yes/no; 0/1) regressed on measured baseline patient covariates. The selection of the covariates used in the development of the propensity score requires careful consideration as the variance of the estimated exposure effect can be strongly influenced by the selection of the variables and how such variables are categorized and made to interact with each other (Brookhart *et al.*, 2006).

The process of selection of variables to include in the propensity score model is a topic of some debate (Austin *et al.*, 2007). Early research suggested using variables associated specifically with treatment allocation (Rosenbaum & Rubin, 1983), while authors have more recently suggested that all variables associated with both treatment allocation and outcome should be included in the propensity score model (Perkins *et al.*, 2000). Rubin suggested that a broad range of variables should be included when building the propensity score. He did, however, warn against the use of variables 'that are effectively known to have no possible connection to the outcomes, such as random numbers, or five-way interactions, or the weather half-way around the world' (Rubin, 2007; p.29). Simulations have demonstrated that the inclusion of variables associated

with the treatment allocation but not the outcome can increase the variance of the treatment effect (Brookhart *et al.*, 2006) and can also reduce the ability to carry out matching based on the propensity score (Austin *et al.*, 2007).

Unlike traditional regression modelling, the propensity score model does not need to be parsimonious and can include numerous covariates, interactions, and nonlinear terms. The propensity score model should be estimated using the structured, iterative process described Rosenbaum and Rubin (1984). First, the initial propensity score should be estimated using the variables as main effects that were identified for inclusion in the model. Balance between these variables should then be assessed. Each potential confounder is regressed on treatment, propensity score quintile, and the treatment and quintile interaction using analysis of variance. If there is evidence of imbalance the propensity score model should be modified by using the square of the variable and through the addition of interactions with other clinically important variables. Higher order terms can be added to further refine the model. This iterative process is repeated until balance is achieved between the potential confounders or when repeating the process is no longer practical. Indeed, the liberal use of interactions and transformations is encouraged in order to create balance between the two treatment groups and therefore better adjust for bias (Shah *et al.*, 2005).

8.2.2 Common support of the propensity score model

The distribution of the estimated propensity score for treated and untreated patients should be plotted to allow for the examination of overlap in the propensity scores of the two treatment groups (Figure 8.1). This can, for example, identify untreated patients with very low propensity scores for whom there would be no treated patients

with whom to match on the propensity score. These patients may not be comparable and their inclusion may lead to a biased estimate of the treatment effect (Rosenbaum & Rubin, 1985; Heckman *et al.*, 1998). This bias can be reduced by enforcing common support, that is, by restricting the analysis to the range of propensity scores that overlap. Enforcing common support excludes untreated patients with a propensity score lower than the lowest propensity score from the treated patients, as well as excluding treated patients whose propensity score are higher than the highest value in the untreated patients.

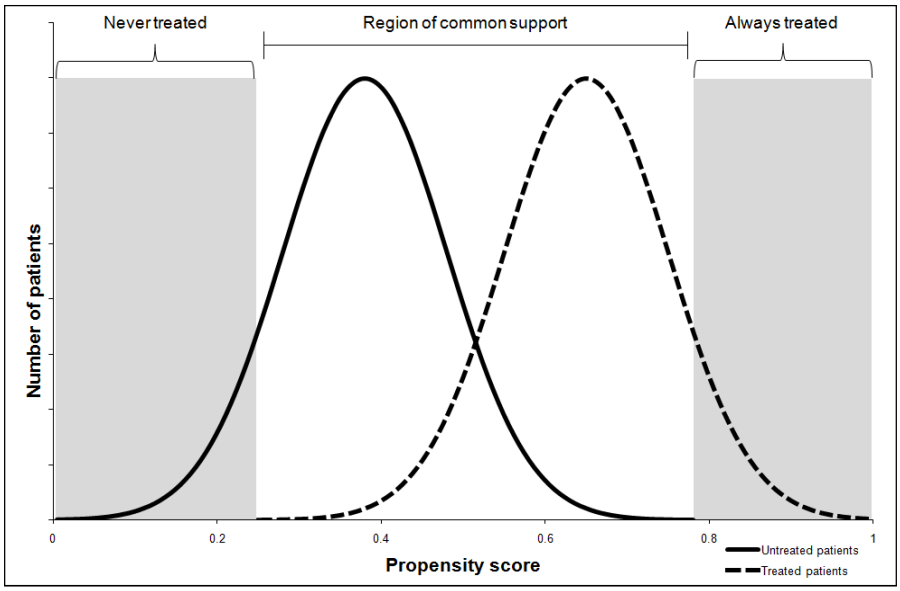


Figure 8.1 – The propensity score distribution demonstrating the area of overlap between treated and untreated patients. Patients with very low propensity scores are never treated and patients with very high propensity scores are always treated (adapted from Schneeweiss & Avorn, 2005).

8.2.3 Propensity score methods

Once estimated, the propensity score can then be used in several ways in order to control for confounding. Some of the commonly used methods are described below.

8.2.3.1 Covariate adjustment using the propensity score

Propensity score can be used as a covariate, in addition to the treatment variable, in a regression model to adjust for the estimate of the treatment effect. Both treatment and the propensity scores are regarded as independent variables in the analysis and it is assumed that the relationship between the propensity score and the outcome is linear. Additional covariates may be included in the regression model.

8.2.3.2 Matching on the propensity score

Matching on the propensity score is an intuitive method to adjust for confounding that does not impose linearity restrictions on the relationship between an outcome and the patient covariates (Zhou, 2005). Matching treated and untreated patients on the propensity score is most commonly carried out through nearest neighbour matching, where an algorithm matches treated patients with the nearest untreated patient based on their similar propensity scores. The majority of studies use one-to-one matching, although one-to-many and many-to-one matches can be made. There is a risk of poor matches using nearest neighbour matching if the nearest neighbour is numerically distant.

Two types of nearest neighbour matching algorithms exist. The more commonly used is the greedy matching algorithm, which randomly selects a treated patient and matches them with the nearest untreated patient, even if that untreated patient could serve as a better match for another treated patient. The other type of algorithm is optimal matching, where the algorithm finds pairs with the smallest average absolute distance between each pair. Pairs that were previously formed can be unmatched if

another pair would have a smaller distance. Optimal matching is difficult to achieve in large data sets for computational reasons (Austin, 2009a).

Greedy nearest-neighbour matching is most often undertaken within a given interval of the treated patient's propensity score (the caliper width) (Austin, 2009c). A treated patient is matched with the nearest untreated patient within a fixed caliper width, which is used to reduce the risk of bad matches by imposing a maximum numerical distance between matches. Deciding on the appropriate caliper width *a priori* is difficult and various calipers have been used in the medical literature. Some authors have carried out matching on the logit of the propensity score, using caliper widths of 0.6 times (Ayanian *et al.*, 2002) and 0.2 times (Austin *et al.*, 2007) the standard deviation of the logit of the propensity score. Other authors have matched on the propensity score using caliper widths ranging from 0.005 (Cole *et al.*, 2002); to 0.01 (Seeger *et al.*, 2003); to 0.1 (Moss *et al.*, 2003). Cochran and Rubin (1973) demonstrated that matching using calipers of width of 0.2 of the standard deviation of the logit of the propensity score eliminated between 98 and 99% of the bias of the confounding variable, assuming that this variable was normally distributed.

A more recent Monte Carlo simulation by Austin (2009c) examined the performance of matching methods using different caliper widths. Using a caliper width of 0.6 times the standard deviation of the logit of the propensity score resulted in the greatest number of matches but also generated estimates with the greatest bias. Austin recommended the use of calipers with width of 0.2 standard deviation of the logit of the propensity score or using calipers of width 0.02 or 0.03 on the propensity score scale.

Nearest neighbour matching on the propensity score can be undertaken with or without replacement. Matching without replacement is most commonly used, where an untreated patient who has been matched with a treated patient is removed from the matching pool and can no longer act as a potential match for other treated patients. Poor matching can occur when there is limited overlap in the distribution of propensity scores of the treated and untreated patients. In matching with replacement, each treated patient is matched with the closest untreated patient, which allows more exposed patients to be used and improves the performance of the match (Lunt *et al.*, 2009). This method does, however, increase the potential for a small number of untreated patients to be used multiple times and can therefore influence the results significantly.

The various matching techniques and algorithms represent a trade-off between bias and efficiency and careful consideration must be given to their application (Table 8.1).

Table 8.1 – The trade-off between decreased bias and improved efficiency in matching techniques (adapted from Hebert, 2010)

Matching techniques	Bias	Inefficiency
1-to-1 matching	↓	↑
1-to-N matching	↑	↓
Matching with replacement	↓	↑
Matching without replacement	↑	↓
Matching with calipers	↓	↑
Matching without calipers	↑	↓

Nearest neighbour matching can also be undertaken using five-to-one (5→1) digit matching. This method matches treated and untreated patients on the first five digits of the propensity score. If this cannot be achieved for a treated patient, then matching is attempted on the first four digits. If a suitable match cannot be found for a treated patient, attempts to find suitable matches on the first three, first two, and finally the first digit of the propensity score are then made.

8.2.3.3 Stratification on the propensity score

In this method, patients are grouped into strata based upon the propensity score and the exposed and unexposed patients are compared within each stratum. The strata are usually equal sized groups created using the quintiles or deciles of the estimated propensity score. Within each stratum of the propensity score the treated and untreated patients are more likely to be directly comparable, similar to a randomized trial (Weitzen *et al.*, 2005). Stratification based on the quintiles of a propensity score has been shown to decrease 90% of bias due to imbalance of potential confounders (Rosenbaum & Rubin, 1984). However, residual imbalance between treated and untreated patients in the upper and lower strata has been demonstrated (Austin *et al.*, 2007), and stratification can lead to greater bias when compared with matching on the estimated propensity score (Austin, 2008d).

8.2.3.4 Weighting by the inverse of the propensity score

Weighting is another method of using propensity scores to estimate treatment effects, although it is used less frequently than other methods. In propensity score weighting, the treated and control observations are re-weighted in order to make them more representative of the population. The weight of a treated patient is defined as the

inverse of its propensity score, while the weight of an untreated patient is defined as the inverse of one minus its propensity score. This method does have potential for serious bias as the procedure can give a high weight to a small number of patients who may not be representative of the population as a whole.

8.2.3.5 Other propensity score and similar methods

Another method that has been suggested but is used infrequently is matching on the Mahalanobis distance between a treated patient and an untreated patient. The Mahalanobis distance $d(i,j)$ is defined as

$$d(i,j) = (u_i - v_j)^T C^{-1} (u_i - v_j) \tag{8.1}$$

where u and v are values of the matching variables for treated patient i and untreated patient j, and C is the sample covariance matrix of the matching variables from the full set of control subjects (D'Agostino, 1998). The Mahalanobis distance scales the distance in variance based on the covariance matrix so that if , for example, the variance of u is twice the variance of v, then an observation needs to be twice as far in order to be equidistant in the Mahalanobis distance (Posner & Ash, 2006). Patients are randomly ordered and the Mahalanobis distance is calculated between the first treated patient and all of the untreated patients. The untreated patient with the smallest Mahalanobis distance is chosen as a match and both the treated and untreated patients are removed from the selection pool. The process is repeated until matches are found for all treated patients. Mahalanobis metric matching can also be carried out including the estimated propensity score and where the logit of the estimated propensity score is used in the calculation of the Mahalanobis distance. Finally, nearest available Mahalanobis metric matching can also be implemented within set calipers that are defined by the estimated propensity score.

Finally, kernel weighting can be used to match every treated patient with the weighted average of the untreated patients. The weights are inversely proportional to the distance between the treated and untreated patients' propensity scores (Baser, 2006). This method results in lower variance as all of the treated patients contribute to the weights.

8.2.4 A comparison of propensity score methods

There is an inherent relationship among the different propensity score methods and they often can generate similar results. For example, matching on the propensity score can be likened to stratification where one is using very small stratum. However, each method has benefits and limitations, which must be considered and acknowledged.

Covariate adjustment using the propensity score is a useful method because no data are lost due to poor overlap or lack of common support between treated and untreated patients, which increases the generalizability of the results. However, this method can be significantly affected by errors in estimation of the propensity score. Regression using the estimated propensity score also assumes a linear relationship between the propensity score and the outcome, which is not assumed with matching or stratifying on the propensity score.

The method of weighting by the inverse of the propensity score is less frequently used and is limited because it attaches more importance to observations at the tail-end of the propensity score distributions. This has potential for serious bias as a small number of patients can be weighted heavily who may not be representative of the entire study cohort. Furthermore, if there is an error in the measurement of the

outcome, treated patients will tend to have smaller propensity scores or treated patients will have larger propensity scores, which will lead to an estimated treatment effect with very large variance.

Several studies have used multiple propensity score methods with the specific goal of making comparisons among the different methods and traditional regression models. Stürmer and colleagues (2005) used various methods including stratification and matching on the propensity score, covariate adjustment using the propensity score, weighting by the inverse of the propensity score and a traditional multivariable model to investigate the effect of non-steroidal anti-inflammatory drug use on mortality. All of the methods produced estimates of the treatment effect that were comparable with the results obtained from prior randomized trials.

A second study examined the effect of statin therapy on mortality following acute myocardial infarction using various propensity score methods, all of which gave similar estimates of treatment effect (Austin & Mamdani, 2006). Covariate adjustment including the propensity score had the lowest variance, but the estimate produced by matching on the propensity score was identical to one obtained from a meta-analysis of trials. The authors also demonstrated that matching and stratifying on the propensity score both improved balance in the baseline characteristics between treated and untreated patients, although greater balance was achieved by matching. This study illustrated the trade-off between matching and stratification on the propensity score: matching improves balance and reduces bias, but may result in a reduction in the sample size if a match cannot be found, which can decrease the precision of the results. Stratification on the propensity score may result in greater bias due to the

strata being heterogeneous in the propensity score—referred to as residual confounding—but results in estimates with greater precision (Austin *et al.*, 2007; Austin, 2008d).

Austin (2009b) used an empirical case study and a simulation study to compare the four main propensity score methods. He demonstrated that matching on the propensity score removed almost all of the systematic differences between treated and untreated patients. Weighting using the inverse of the propensity score was comparable to matching, while stratification on the propensity score and covariate adjustment using the propensity score resulted in greater residual differences in baseline characteristics between treated and untreated patients.

There is some evidence that the various propensity score methods can perform differently depending on which method is used and the context in which it is used. Kurth and colleagues (2006) analyzed the effect of tissue plasminogen activator on mortality and demonstrated that the different propensity score methods can have significantly divergent results. They found that covariate adjustment and stratification on the propensity score tended to overestimate the treatment effect while weighting by the inverse of the propensity score overestimated the treatment effect by a factor of approximately 10 because of many very small propensity score values. The differences were most likely due to the inclusion of patients with low propensity scores who were uncommon in the treated group, common in the untreated group, and significantly different from the rest of the patients in their risk of death. This was demonstrated when the analyses were limited to patients with a propensity score of greater than 0.05; all of the propensity score methods gave fairly similar results. The

authors argued that the variation in the results obtained by the different methods does not suggest that one of the methods is better for controlling confounding. Instead, one must investigate the potential causes of the divergent results.

The different propensity score methods have their advantages and disadvantages and can produce different results, which may be valid but depend on the research question implied by the adjustment method. Matching is often chosen because it can generate an easier, more transparent analysis, but this method is limited by the loss in the sample size due to incomplete matching. Stratification on the propensity score improves precision but may result in greater bias due heterogeneous strata. Weighting by the inverse of the propensity score is not very often used, primarily because in cases with extreme propensity score values, the variance of the weighted estimator can be very large. Generally, it is not apparent whether any of various propensity score methods is superior to the others. Neither is it certain whether propensity score methods are preferable to traditional regression models.

8.2.5 Diagnostics of propensity score matching methods

Balance diagnostic tests determine whether the propensity score model has been adequately specified by identifying whether balance has been achieved between the treated and untreated patients with respect to the covariates used to define the propensity score (Austin, 2009a). Balance diagnostics differ depending on how the propensity score was used to adjust for confounding, although all of the diagnostics use a method based on the standardized difference to compare the balance in measured baseline covariates between treated and untreated patients. The use of standardized differences is preferable to significance testing and probability values

because they are not influenced by sample size and are less sensitive to bias. The standardized difference (*d*) for continuous variables is defined as

$$d = \frac{100(\overline{x}_{treated} - \overline{x}_{untreated})}{\sqrt{\frac{s^2_{treated} + s^2_{untreated}}{2}}} \tag{8.2}$$

where $\bar{x}$ represents the sample mean and *s* the sample variance of the covariate in treated and untreated patients. For dichotomous variables, the standardized difference (*d*) is defined as

$$d = \frac{100(\hat{p}_{treated} - \hat{p}_{untreated})}{\sqrt{\frac{\hat{p}_{treated}(1-\hat{p}_{treated}) + \hat{p}_{untreated}(1-\hat{p}_{untreated})}{2}}} \tag{8.3}$$

where $\hat{p}$ is the prevalence of the dichotomous variable in treated and untreated patients. The standardized difference allows for the comparison of the relative balance between different covariates measured in different units (Austin, 2008e). A standardized difference of 10 percent or greater has been demonstrated to indicate significant imbalance in the patient characteristic (Normand *et al.*, 2001).

In addition to the standardized differences of the potential confounders in the matched sample, means of continuous covariates and the frequency distribution of categorical variables between treated and untreated patients should be reported (Austin, 2009a). The adequacy of matching can further be assessed by calculating the percent bias reduction, which is defined as

$$\%\ bias\ reduction = 1 - \left(\frac{(|d_{matched}|)}{(|d_{unmatched}|)} \right) \tag{8.4}$$

where *d* is the standardized difference obtained in either equation (8.2) or (8.3) .

When many-to-one matching has been carried out, a weighted standardized difference should be used, using the inverse of the number of untreated subjects within a given matched set as the weight for each untreated patient (Austin, 2008a). When stratification on the propensity score is used, within-strata standardized differences are used to compare the distribution of baseline confounders between treated and untreated patients within the same propensity score stratum. A mean standardized difference can then be combined across the strata (Austin, 2009b).

Two balance diagnostics have been suggested when covariate adjustment with the propensity score is used. A weighted conditional standardized difference can be computed, which compares the conditional difference in baseline covariates between the two treatment groups. A qualitative method using a quantile regression model conditional on the propensity score can compare the distribution of continuous baseline covariates between treated and untreated patients (Austin, 2008c).

Some authors have used the Hosmer–Lemeshow test to assess the goodness of fit (GoF) of the propensity score model. The GoF assesses how well the model describes the data and whether differences between observed values from the data and the values predicted from the model are small and random. Lack-of-fit may occur due to the misspecification of continuous variables, the use of inappropriate interaction terms, or through the omission of important confounders (Hosmer *et al.*, 1997).

Another diagnostic test that is often presented is the area under the receiver operator characteristic (ROC) curve (also called the *c*-statistic). The *c*-statistic assesses the discrimination of the propensity score model, or how well the estimated propensity

score classifies patients into treated or untreated groups. A value of 0.5 would suggest that the propensity score model was as effective as tossing a coin, compared with a value of 1.0 that suggested that the model perfectly predicted whether a patient was treated or untreated (Weitzen *et al.*, 2005).

Although both the GoF test and *c*-statistic are often reported (Weitzen *et al.*, 2004), several authors have warned against the use of these tests of model discrimination and specification. Previous simulation studies have demonstrated that both tests of GoF and the *c*-statistic fail to identify whether an important confounding variable has been omitted from the propensity score model or when a model is misspecified (Weitzen *et al.*, 2005; Austin *et al.*, 2007; Austin, 2009a). Indeed, no association has been demonstrated between the *c*-statistic and the ability of the propensity score model to balance potential confounders between treated and untreated patients (Austin *et al.*, 2007). A propensity score with a high *c*-statistic (e.g. 0.90) can have significant levels of non-overlap in the distributions of propensity scores in treated and untreated patients, which can lead to significant bias in the estimation of the treatment effect.

Other statistical tests, in addition to the calculation of standardized differences, have also been suggested to ensure that balance has been achieved between the two treatment groups on potential confounding factors in matched cohorts. The variance of variables can also be compared between treated and untreated patients (Imai *et al.*, 2008; Austin, 2009a). Austin demonstrated in a simulation study that incorrectly specified propensity models can be identified by comparing the ratios of variances (Austin, 2009a). Ho and colleagues (2007) and Austin (2009a) have both suggested comparing the means of interactions of pairs between treated and untreated patients,

which can determine if covariances are similar between the two treatment groups. Austin's simulation study demonstrated that a large standardized difference indicated an incorrectly specified propensity score model.

Graphical methods have also been suggested as a way to assess balance, including the use of quantile-quantile plots and side-by-side box plots (Ho *et al.*, 2007; Imai *et al.*, 2008). These graphs can be used to compare the distribution of baseline covariates between treated and untreated groups, and help to provide additional evidence that the propensity model has been correctly specified.

8.2.6 The use of propensity score methods

Although propensity score methods were first fully described in 1983 (Rosenbaum & Rubin) as a tool to reduce confounding, it has only been in recent years where an increase in their use has been observed (Austin, 2008b). Indeed, the National Library of Medicine did not introduce the medical subject heading (MeSH) 'propensity score' until July 2009.

A search of Medline (via OVID) on 1 January 2011 using the text words propensity score$ or propensity match$ or the propensity score MeSH term identified 2279 publications. A dramatic increase in the number of publications and the proportion of Medline citations referencing propensity score methods over time is evident (Figure 8.2).

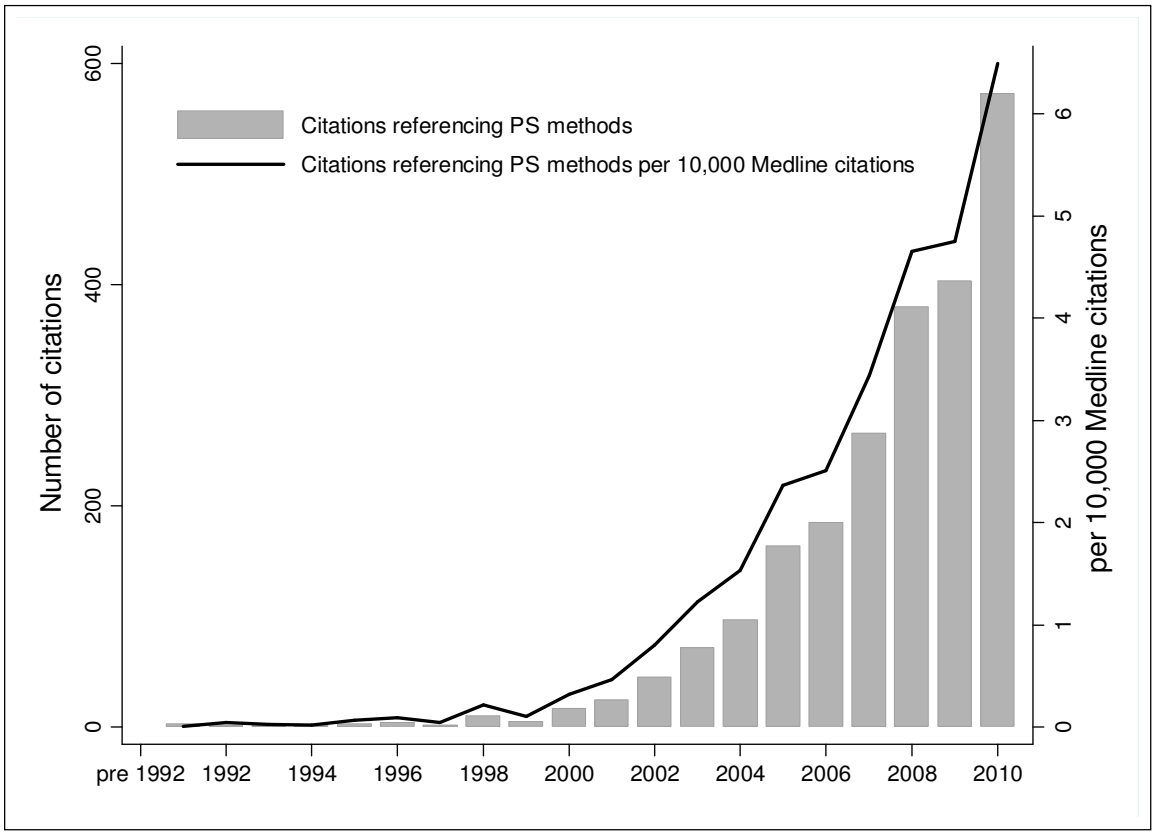


Figure 8.2 – Total number of citations and number of citations per 10,000 publications in Medline referencing propensity score (PS) methods by year of publication

Several systematic reviews have been carried out over the past six years examining the nature and quality of medical studies using propensity score methods (Table 8.2). An early systematic review by Weitzen and colleagues (2004) examined 47 studies published in 2001 that used propensity score methods. The majority of studies (25) adjusted the model using the propensity score, 7 used matching based on the estimated propensity score, and 13 studies stratified on the propensity score. Two studies provided no clear description of the method used.

Shah and colleagues (2005) reviewed 43 studies that published results using both traditional regression and propensity score methods to control for confounding. The majority of studies (22) used the propensity score as a covariate in a regression model,

13 studies matched on the propensity score, and 8 studies used stratification on the propensity score. Fewer than half (19) of the studies reported whether balance had been achieved in the confounders used to estimate the propensity score. The estimates generated by propensity score methods tended to be more conservative than those from traditional regression analyses.

A large review examined 177 studies published until the end of 2003 that used propensity score methods (Stürmer *et al.*, 2006a). Fifty-one studies used matching on the propensity score. Of the 69 studies that presented the results of both propensity score methods and traditional methods, only 9 (13%) had an effect estimate that differed by more than 20%. Seventy-three studies (41%) reported the discrimination of the propensity score model using the area under the ROC curve, which ranged from 0.56 to 0.93.

Table 8.2 – A summary of systematic reviews of published studies that have used propensity score methods

Study	Number of studies	Inclusion criteria	Number of covariates in PS model	PS methods used, n (%)	If matching on PS used		
					Balance checked	% treated matched	Method of matching described
Weitzen (2004)	47	Studies published in 2001 that used PS methods	Mean: 19 Range: 3–102 NR: 13 (28%)	25 (53%) – covariate adjustment using PS 13 (28%) – stratified on PS 7 (15%) – matching on PS 2 (4%) – no description of PS methods	7/7 (100%)	NR	NR
Shah (2005)	43	Studies that used both traditional regression models and PS methods to control for confounding	NR	22 (51%) – covariate adjustment using PS 11 (30%) – matching on PS 10 (19%) – stratified on PS	8/11 (73%)	NR	NR
Stürmer (2006a)	177	Studies published until 31 December 2003 that used PS methods	Mean: 21 Range: 2–112 NR: 29 (16%)	87 (49%) – multivariable model with PS 81 (46%) – multivariable model without PS 69 (39%) – covariate adjustment using PS 51 (29%) – matching on PS	NR	Median: 90 Range: 26–100 NR: 2	NR
Austin (2008b)	47	Studies published between 1996 and 2003 that used matching on the PS	NR	47 (100%) – matching on PS	39/47 (83%)	NR	18 (38%) – matching using various caliper widths 15 (32%) – NR 10 (21%) – 5→1 digit matching

Study	Number of studies	Inclusion criteria	Number of covariates in PS model	PS methods used, n (%)	If matching on PS used		
					Balance checked	% treated matched	Method of matching described
							3 (6%) – nearest neighbour matching† 1 (2%) – PS quintiles
Austin (2007)	60	Studies that used matching on the PS published from 1 January 2004 to 31 December 2006 from 5 cardiovascular surgery journals	NR	60 (100%) – matching on the PS	49 (82%)	NR	31 (52%) – NR 20 (33%) – 5→1 digit matching 9 (2%) – matching using various caliper widths
Austin (2008e)	44	Studies that used matching on the PS published from 1 January 2004 to 31 December 2006 from 7 general cardiology journals	NR	44 (100%) – matching on the PS	30 (68%)	NR	20 (45%) – NR 13 (30%) – matching using various caliper widths 11 (25%) – 5→1 digit matching
Gayat (2010)	47	Studies published from 2006 to 2009 from 7 intensive care and anaesthesiology journals	Median: 15 Range: (9–22)	26 (55%) – matching on PS 9 (19%) – stratification on PS 12 (26%) – covariate adjustment using the PS 0 – weighting on PS	20/26 (77%)	32 (24–54)	13 (50%) – 5→1 digit matching 8 (31%) – nearest neighbour matching† 5 (19%) – NR

NR = not reported; PS = propensity score; † = no description of whether a caliper width was used with nearest neighbour matching

A systematic review examined studies that were published in the medical literature between 1996 and 2003 that carried out matching on the propensity score (Austin, 2008b). Forty-seven studies were identified, of whom 15 (32%) did not report how matched pairs were formed. Eight studies did not assess whether matching on the propensity score resulted in balance between potential confounders in the treated and untreated patients. Analytic methods to estimate the treatment effect did not take into account the matched nature of the cohort in 26 studies.

Another systematic review had a more narrow scope and examined studies published from 2004 to 2006 from the cardiovascular surgery literature (Austin, 2007). Austin identified 60 studies that carried out matching using an estimated propensity score. Almost 30% (17) of the papers did not report how the matched pairs were created. There was poor reporting of whether balance on the potential confounders had been achieved between treated and untreated patients, as 18% of the studies did not report this information. Over two-thirds (39) of the studies used statistical methods to estimate the treatment effect that did not take into account the matched nature of the cohort.

Austin (2008e) reported a second similar systematic review of studies undertaking propensity score matching published within the same time frame (2004 to 2006), but which were reported in the cardiology literature. There was no overlap in the studies from his other review. Austin identified 44 studies of which 24 (55%) reported how matches were formed. Fourteen studies did not assess balance in the confounders between treated and untreated patients. Almost half of the studies (21) used statistical

techniques to assess the statistical significance of the treatment effect that were inappropriate given the matched cohort.

The most recent systematic review was carried out by Gayat and colleagues (2010), who examined studies published in intensive care and anaesthesiology journals from 2006 to 2009. The average number of covariates used to estimate the propensity score was 15 (range 9–22) in the 47 studies reviewed. The majority of studies (26) carried out matching on the propensity score, of which 23 (79%) presented data on balance between confounders. Nine articles used stratification of the propensity score and 12 used the propensity score in a regression model. No studies used weighting on the propensity score.

The seven systematic reviews identified several common themes. First, reviewers called for greater clarity and transparency in how the propensity score was estimated, with a clearer description of covariate selection and how these covariates were chosen. Secondly, when propensity score matching techniques are used, authors should explicitly state how the pairs were created and whether sampling was carried out with or without replacement. Thirdly, the distribution of baseline characteristics should be compared. This should be carried out using standardized differences and most authors discouraged the use of GoF tests and the *c*-statistic. Weitzen and colleagues (2004) did not explicitly discourage the use of these tests of goodness of fit and model discrimination but later the same authors published a study demonstrating that these tests failed to detect missing confounders in a propensity score model (Weitzen *et al.*, 2005). Finally, appropriate statistical methods that take into the

account the lack of independence due to matching should be used when analyzing a matched cohort.

The most recently published systematic review by Gayat and colleagues (2010) described an improvement in the use of more robust methods to assess balance between confounders (e.g. the use of standardized differences) and a general turn towards the use of matching as the primary propensity score technique. The authors did, however, describe continued failings in how the propensity score was estimated and in the statistical methods used to determine the treatment effect.

The use of propensity score methods is not restricted to the medical literature. These methods have also been used in observational research investigating other 'treatments' that have been non-randomly allocated such as the exposure to different educational (Yanovitzky *et al.*, 2005; Zanutto *et al.*, 2005) or economic programmes (Dehejia & Wahba, 2002).

8.2.7 Benefits and limitations of propensity score methods

Propensity score methods have been shown to be robust and although they may slightly decrease the strength of the association between treatment and outcome, they do not tend to give significantly different results than those obtained from traditional regression models (Shah *et al.*, 2005; Stürmer *et al.*, 2006a). They may also provide additional benefit by restricting the analysis to groups that overlap with respect to potential confounders and therefore are more readily comparable (Hill *et al.*, 2004). The lack of sufficient overlap cannot be detected in traditional multivariable modelling.

Another advantage of propensity score methods is that they have been shown to perform particularly well when there are few events per potential confounding variable. When an outcome is rare and the drug treatment is relatively common—such as rhabdomyolysis with statins—there may be too few data with which to model the relationship between an outcome and patient covariates. For example, if several thousand patients are assigned to one of two drug treatments and only thirty outcomes are observed, standard logistic regression with multiple covariates would not be possible as ten events per covariate is usually considered to be the minimum requirement for stable, unbiased estimates (Peduzzi *et al.*, 1996; Stürmer *et al.*, 2006a). When treatment is common, there are sufficient data with which to model the relationship between treatment and patient outcomes and therefore calculate a propensity score. The use of a single propensity score, which has been estimated with many patient covariates, can therefore control simultaneously for many measured covariates, even when there are too many covariates to model their relationships with the rare outcome (Braitman & Rosenbaum, 2002).

A simulation study by Cepeda and colleagues (2003) demonstrated that when there were seven or fewer events per confounder, propensity score estimates were less biased, more robust and more precise than logistic regression methods. This has important implications for pharmacoepidemiological research, where studies often examine rare outcomes (Glynn *et al.*, 2006).

There is also some evidence that stratification on the propensity score may have some benefit over traditional methods when effect modification is present, because a summary measure is calculated over the strata. This has particular relevance to

pharmacoepidemiology where one can infer the relationship between drug and an outcome for an entire treated population (Stürmer *et al.*, 2006b).

Propensity score methods can adjust for selection bias in observational studies and can perform well when modelling rare outcomes. They are not, however, 'magic bullets' that can eliminate all bias in observational studies (Shah *et al.*, 2005). Some authors have noted the 'perceived opacity of the statistical process', which can cause propensity score methods to have a 'very black box feeling'(Nuttall & Houle, 2008).

Propensity score methods cannot control for confounding by unmeasured factors, and it is these unmeasured confounders that can still lead to biased results. This is the main limitation of observational research because, compared with randomized studies, the lack of randomization cannot balance the distribution of all covariates, observed or unobserved. Residual confounding bias cannot be excluded, particularly when one is carrying out research in databases where information on confounders is limited. Although sensitivity analyses techniques have been suggested as methods to quantitatively assess the degree of residual confounding, these methods are not often used (Schneeweiss, 2006).

There can also be errors in the model used to estimate the propensity score, and these can result in bias. Although the model need not be parsimonious, the addition of extraneous variables into the model should be avoided as they may not bias the results but can increase the variance (Heckman *et al.*, 1998).

Propensity score methods also work better in large samples as substantial imbalance has been demonstrated in small samples due to an increase in the variance of estimated effects (Rubin, 1997). The study sample size needs to be sufficiently large, particularly if stratification is used, in order to ensure that there are sufficient numbers in each stratum. Comparison between treatment groups may be impossible if a stratum only contains patients that belong exclusively to one treatment group. A large sample size is also important if one uses the propensity score to match treated with untreated patients, as there is potential for a significant decrease in the patient population due to incomplete matching, particularly when common support—restricting the analysis to the range of propensity scores that overlap—is enforced.

When working with propensity scores, the distribution of the propensity scores between the treated and untreated patients should overlap. A significant lack of overlap, or propensity score distributions that are very dissimilar, may introduce significant error.

Finally, propensity score methods can be limited by missing data values. Because propensity scores are estimated using a wide range of covariates, many patients are likely to be excluded from the analysis as a significant proportion of patients often have missing values for at least one covariate.

8.3 Application of propensity score methods to the GPRD

8.3.1 Introduction

Biochemical testing of patients treated with antihypertensive drugs is not randomly allocated and therefore there are systematic differences between patients who are tested and those who are not. This may confound any relationship between testing and adverse patient outcomes. I will therefore use both propensity score methods and multivariable regression modelling to adjust for potential bias due to the non-random nature of baseline biochemical testing. Some of the adverse outcomes are rare and the propensity score methods may generate estimates that are more precise and less biased than the traditional methods when the number of events per confounder is low.

The aims of this analysis were:

1. To examine the relationship between baseline biochemical testing and adverse patient outcomes;
2. To use propensity score methods to adjust for potential confounding in baseline testing;
3. To make comparisons between the results obtained from propensity score methods and traditional multivariable regression models.

8.3.2 Methods

8.3.2.1 Propensity score development

I wished to develop a propensity score model that determined the probability of each patient having any baseline laboratory test (creatinine, potassium, sodium, or urea) in the six months before the start of antihypertensive treatment. The propensity score was estimated using a logistic regression model where the binary outcome was any baseline biochemical test regressed on measured baseline patient covariates.

The development of the propensity score model was an iterative process (Figure 8.3). Twenty-one patient characteristics that were considered to be clinically relevant and associated with baseline testing or the outcomes *a priori* were used in the development of the initial propensity score. These characteristics include: age, sex, socioeconomic status (SES) score, body mass index (BMI), number of units of alcohol per week, systolic blood pressure, diastolic blood pressure, diabetes, smoking status, hypothyroidism, peripheral vascular disease (PVD), antihypertensive drug treatment, concomitant treatment with carbamazepine, concomitant treatment with non-steroidal anti-inflammatory drugs (NSAID), concomitant treatment with selective serotonin re-uptake inhibitors (SSRI), concomitant treatment with tricyclic antidepressants, concomitant treatment with trimethoprim, year of cohort entry, ten or more prescriptions in the year prior to starting antihypertensive treatment, prior evidence of electrolyte ADRs, and eight or more GP practice visits in the year prior to starting antihypertensive treatment.

Once the initial propensity score was determined, patients were then stratified into quintiles of the estimated propensity score. Each potential confounder was then regressed on baseline testing, propensity score quintile and the baseline testing and quintile interaction using analysis of variance. A P value of <0.05 for either the baseline testing effect or the baseline testing and propensity score quintile interaction suggested imbalance between the patients with and without baseline testing for that specific confounder. When there was evidence of imbalance, interactions with other variables in the model or higher order terms for continuous variables were added.

This process was repeated until there was no evidence of imbalance between the potential confounders or where it was evident that repeating the process was no longer practical. Once the propensity score was determined for each patient, the common support condition was imposed by deleting patients with no baseline testing with lower estimated propensity scores than patients with any baseline testing and deleting patients with baseline testing with higher propensity scores than patients with no baseline testing.

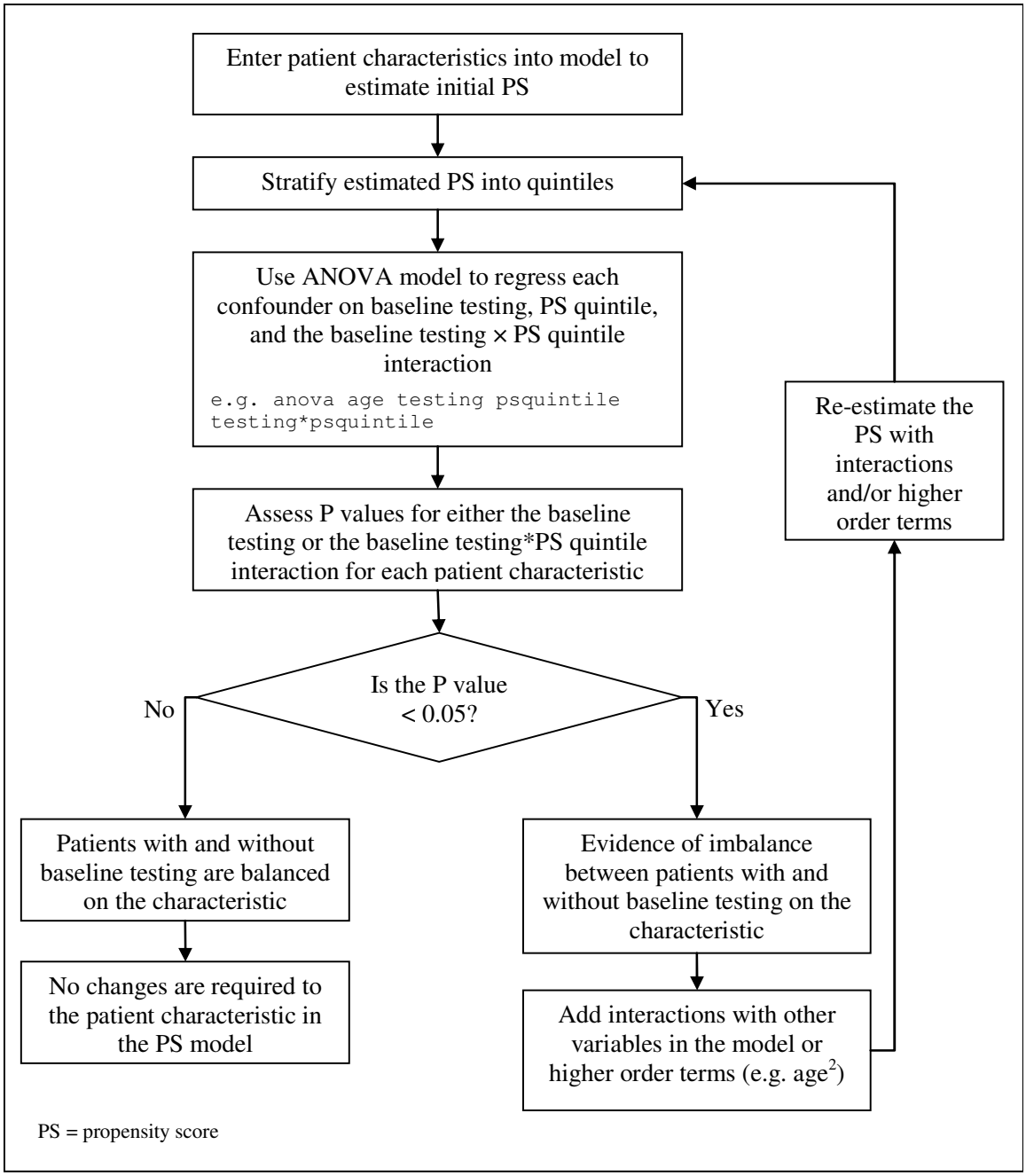


Figure 8.3 – Flow chart describing the estimation of the propensity score

A histogram and a kernel density estimate were used to examine the distribution of the estimated propensity score and make comparisons between patients with and without baseline testing. Two-sample Kolmogorov–Smirnov tests were used to assess differences between the propensity score distributions. Boxplots and quantile-quantile

plots were also used to graphically examine the distribution of the propensity score within quintiles.

8.3.2.2 Assessment of balance between confounders

Differences in patient characteristics between the two groups were examined using two methods: (1) Chi-squared tests were used to compare binary characteristics and t-tests were used for continuous variables; and (2) standardized differences for each patient characteristic. The percentage reduction in bias was also calculated.

8.3.2.3 Relationship between baseline monitoring and patient outcomes

Several methods were used to determine the relationship between baseline monitoring and four patient outcomes within six months of starting antihypertensive treatment: antihypertensive drug discontinuation, hospital admission, death, and any adverse drug reaction. These outcomes were identified using methods described in section 3.3.6.

One-to-one matching using the propensity score

Matched pairs were created using a greedy matching algorithm without replacement, using a caliper of 0.001 of the estimated propensity score. Various caliper widths of 0.01, 0.02, 0.03, 0.1, and 0.2 times the standard deviation of the logit propensity score were also used to assess the impact of different widths on the number of matches and the reduction in bias. The *psmatch2* programme written for Stata® (Leuven & Sianesi, 2003) was used to carry out one-to-one matching. This programme can implement different kinds of matching and has the ability to graph common support (*psgraph*) and carry out covariate imbalance testing (*pstest*). Conditional logistic

regression modelling was then used to estimate the relationship between baseline testing and adverse patient outcomes in the matched cohort of patients.

Adjustment using the propensity score

The estimated propensity score was used as an independent covariate in a univariable logistic regression model.

Stratification using propensity score quintiles

The propensity score quintile was used as covariate in a univariable logistic regression model.

Standard/traditional multivariable regression modelling

Traditional logistic regression modelling was used to estimate the relationship between baseline testing and patient outcomes. Patient characteristics that were statistically significant at the $P<0.05$ level or were biologically plausible were entered into a multivariable model using backwards stepwise variable selection.

The coding used in Stata® for the analyses described above is presented in Appendix 13.

8.3.3 Results

8.3.3.1 Cohort characteristics

I identified 38 600 patients with complete data for all of the covariates who were started on antihypertensive treatment between 2000 and 2003. 17 777 (46.1%) had any baseline testing in the six months prior to the initiation of antihypertensive treatment. The propensity score model included 21 main effects and three two-way interactions (age × sex, age × diabetes, diabetes × sex).

The distribution of patient characteristics between patients with and without baseline testing is presented in Table 8.3. Patients who did have baseline testing were significantly more likely to be male, to have a higher BMI, to drink a greater number of units of alcohol per week, to have diabetes mellitus, to be smokers, to have evidence of PVD, to be started on an ACE inhibitor as part of their antihypertensive treatment, and to have a greater number of GP practice consultations in the year prior to the start of antihypertensive treatment.

Table 8.3 – Patient characteristics at baseline

Characteristics	Any baseline testing (N=17 777)	No baseline testing (N=20 823)	P value
Demographics			
Age	59.7 (11.9)	59.8 (12.4)	0.550
Male sex	10 088 (56.8)	10 985 (52.8)	<0.0005
SES score	22.7 (17.2)	21.2 (15.8)	<0.0005
BMI	28.5 (5.2)	28.2 (5.0)	<0.0005
Number of units of alcohol per week	11.4 (14.3)	10.9 (14.3)	0.002†
Vital signs			
Systolic blood pressure	169.9 (19.0)	170.0 (19.9)	0.497
Diastolic blood pressure	97.5 (11.2)	97.5 (11.1)	0.613
Co-morbidities			
Diabetes	2 293 (12.9)	1 253 (6.0)	<0.0005

Characteristics	Any baseline testing (N=17 777)	No baseline testing (N=20 823)	P value
Smoking	9 177 (51.6)	10 223 (49.1)	<0.0005
Hypothyroidism	92 (0.52)	115 (0.54)	0.641
PVD	140 (0.79)	117 (0.56)	0.007
Antihypertensive treatment			
ACE inhibitor	3 433 (19.3)	2 753 (13.2)	<0.0005
Alpha-blocker	183 (1.0)	211 (1.0)	0.875
AT-II receptor antagonist	363 (2.0)	418 (2.0)	0.810
Beta-blocker	4 180 (23.5)	5 353 (25.7)	<0.0005
Ca-channel blocker	1 176 (6.6)	1 518 (7.3)	0.010
Combination	135 (0.76)	231 (1.1)	<0.0005
Loop diuretic	62 (0.35)	93 (0.45)	0.130
Thiazide diuretic	8 245 (46.4)	10 246 (49.2)	<0.0005
Concomitant medications			
Carbamazepine	75 (0.42)	96 (0.46)	0.564
NSAIDs	2 498 (14.1)	2 806 (13.5)	0.101
SSRIs	565 (3.2)	736 (3.5)	0.053
Tricyclic antidepressants	655 (3.7)	807 (3.9)	0.327
Trimethoprim	297 (1.7)	306 (1.5)	0.112
Year of study entry			
2000	2 820 (15.9)	5 122 (24.6)	<0.0005
2001	3 917 (22.0)	5 202 (25.0)	<0.0005
2002	5 017 (28.2)	5 528 (26.6)	<0.0005
2003	6 023 (33.9)	4 971 (23.9)	<0.0005
10 or more prior prescriptions	8 021 (45.1)	9 298 (44.7)	0.357
Prior electrolyte ADRs	34 (0.19)	29 (0.14)	0.207
8 or more GP consultations	5 329 (30.0)	5 636 (27.1)	<0.0005

* Binary variables are presented as proportions (%), continuous variables are presented as mean (standard deviation); † Mann-Whitney test for non-parametric data

8.3.3.2 Distribution of the estimated propensity score

Imposing common support led to the deletion of 150 patients. Figure 8.4 demonstrates the histogram and the kernel density estimate of the estimated propensity score, which ranged from 0.009 to 0.946. The estimated propensity score distribution in patients with and without baseline testing differed significantly (Kolmogorov–Smirnov P<0.0005). Patients who did have baseline testing tended to have a higher propensity score.

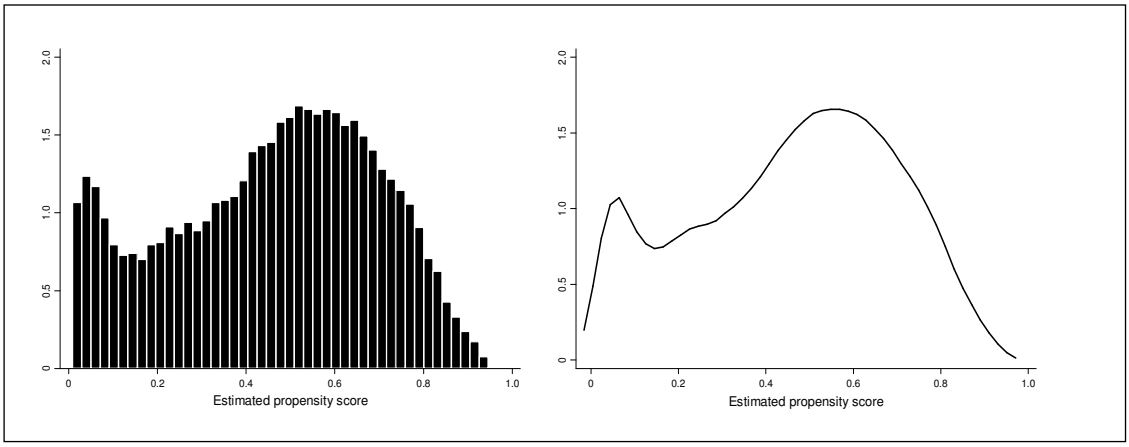


Figure 8.4 – Histogram and kernel density estimate of the estimated propensity score

Figure 8.5 presents box plots of the estimated propensity scores for patients with and without baseline testing within each quintile. The distribution of the propensity scores within each quintile is generally similar. There is, however, some evidence of a significant difference in the median propensity score in quintile 1 between patients with and without baseline testing.

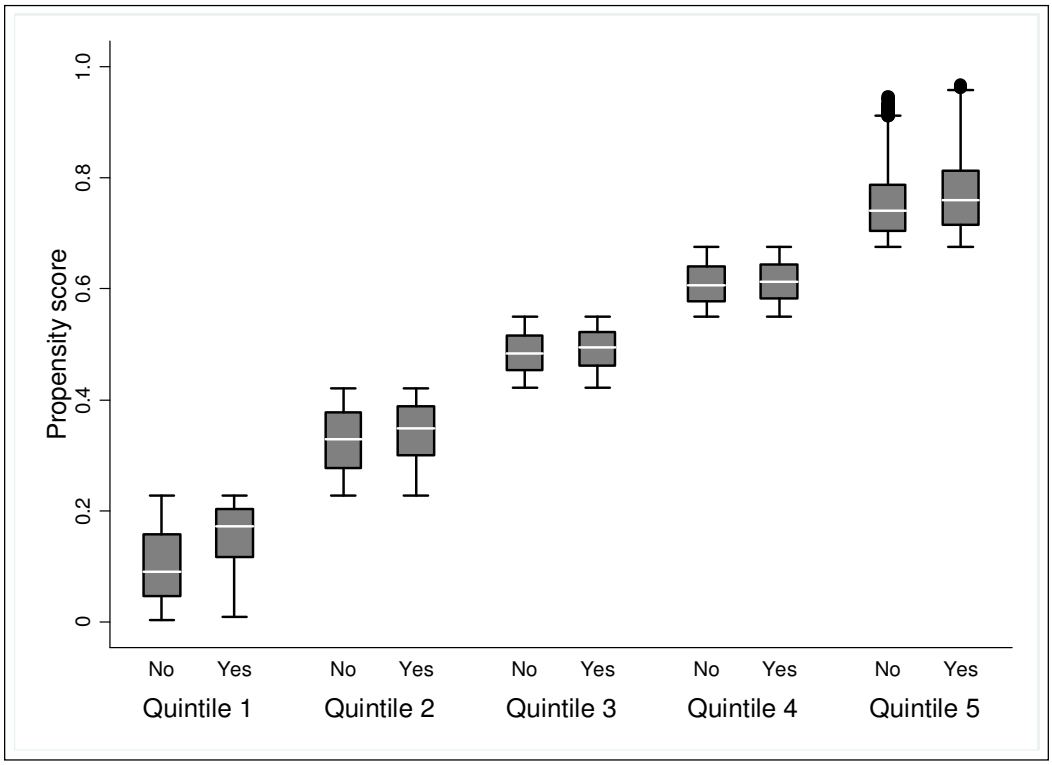


Figure 8.5 – Comparison of propensity score by propensity score quintile and baseline testing

The distribution of the propensity scores was also examined using quantile-quantile plots (Figure 8.6). Again, these plots demonstrated that there is evidence of residual imbalance in the first quintile.

Table 8.4 compares the patients with and without baseline testing by propensity score quintile for each patient characteristic. In general, within each quintile, patients with and without baseline testing are similar. There is some evidence of residual imbalance in patient characteristics within some quintiles, particularly in quintiles 1 and 5. For example, patients with baseline testing had a lower diastolic blood pressure measurement and drank more units of alcohol per week in quintile 1.

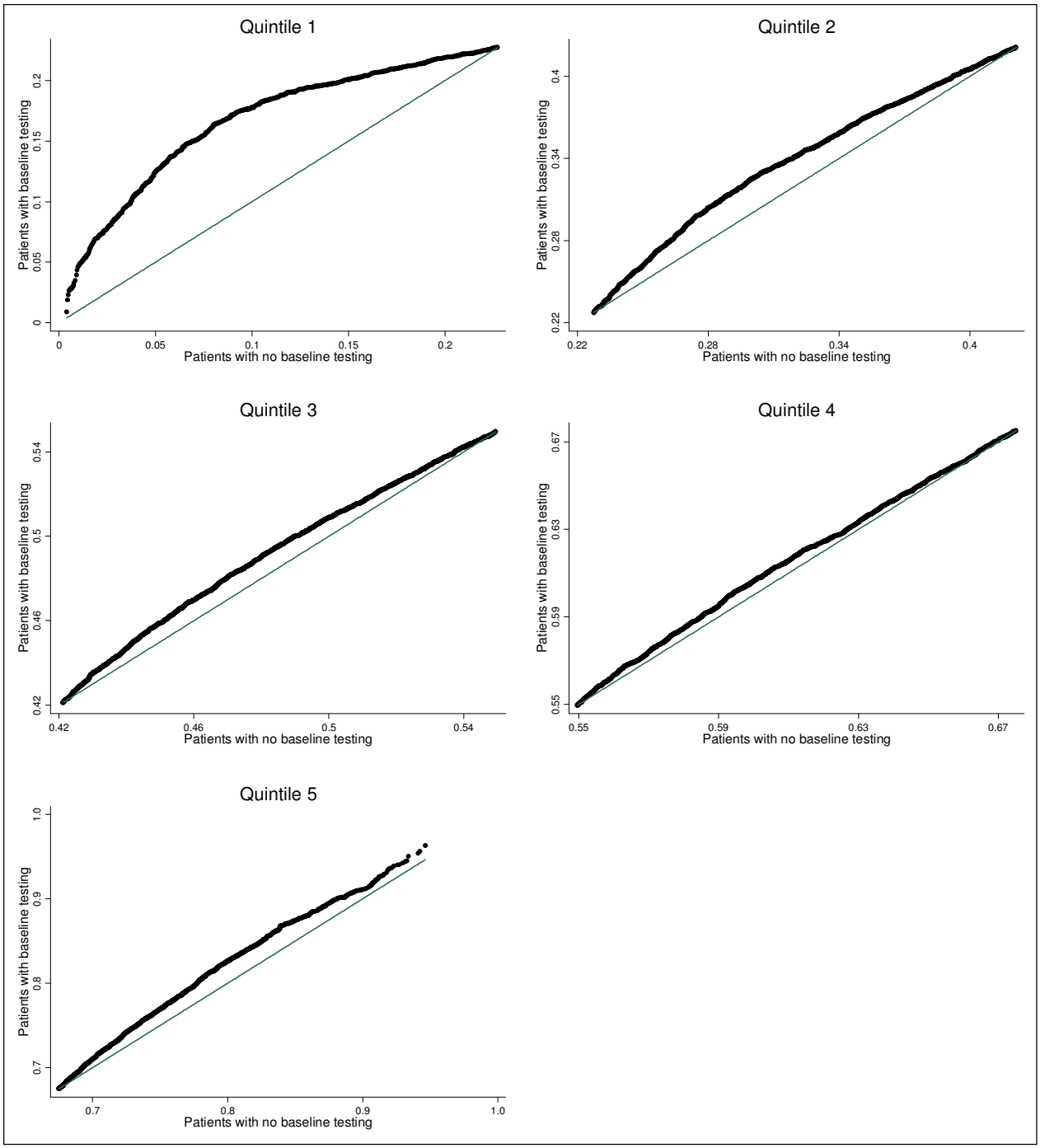


Figure 8.6 – Distribution of propensity score in patients with and without baseline testing by propensity score quintile. The 45° line indicates identical distributions.

Table 8.4 – Comparison of patients with and without baseline testing by propensity score quintile

	Quintile 1			Quintile 2			Quintile 3			Quintile 4			Quintile 5		
	Yes (n=675)	**No (n=7 045)**	**P value**	**Yes (n=2 632)**	**No (n=5 088)**	**P value**	**Yes (n=3 994)**	**No (n=3 726)**	**P value**	**Yes (n=4 841)**	**No (n=2 879)**	**P value**	**Yes (n=5 635)**	**No (n=2 085)**	**P value**
Demographics															
Age	60.0 (12.3)	59.8 (12.3)	0.680	60.0 (12.2)	59.8 (12.5)	0.746	59.6 (11.9)	59.8 (12.5)	0.642	59.9 (11.9)	59.9 (12.3)	0.593	59.5 (11.7)	59.4 (12.3)	0.708
Male sex	356 (52.7)	3575 (508)	0.322	1320 (50.2)	2540 (49.9)	0.848	2128 (53.3)	1969 (52.8)	0.702	2736 (56.5)	1600 (55.6)	0.420	3548 (63.0)	1301 (62.4)	0.648
SES score	20.1 (14.0)	20.0 (14.9)	0.815	20.3 (14.2)	20.5 (14.2)	0.459	22.0 (16.3)	22.6 (16.7)	0.113	22.5 (17.5)	22.3 (17.5)	0.585	24.7 (18.9)	23.3 (18.0)	0.0034
BMI	28.4 (5.3)	28.1 (5.0)	0.077	28.0 (5.0)	28.0 (5.0)	0.597	28.1 (5.0)	28.1 (5.0)	0.951	28.3 (5.1)	28.4 (5.2)	0.497	29.0 (5.3)	28.7 (5.10)	0.008
Number of units of alcohol/week	11.7 (16.1)	10.5 (13.6)	0.027†	10.6 (13.8)	10.7 (13.9)	0.787†	10.8 (13.8)	11.2 (15.6)	0.251†	11.5 (14.2)	11.1 (13.8)	0.250†	11.9 (14.7)	12.0 (15.4)	0.784†
Vital signs															
Systolic blood pressure	169.2 (18.8)	169.5 (19.6)	0.680	170.3 (19.3)	170.2 (20.3)	0.826	170.1 (19.2)	170.6 (20.0)	0.354	170.6 (19.2)	170.7 (19.8)	0.921	168.8 (18.6)	169.0 (19.6)	0.663
Diastolic blood pressure	96.3 (11.4)	97.3 (10.7)	**0.017**	97.2 (10.8)	97.5 (11.1)	0.244	97.9 (10.9)	97.9 (11.2)	0.828	98.1 (11.1)	97.7 (11.6)	0.110	96.9 (11.6)	97.1 (11.9)	0.394
Co-morbidities															
Diabetes	83 (12.3)	268 (3.8)	**<0.0005**	122 (4.6)	183 (3.6)	**0.026**	152 (3.8)	146 (3.9)	0.797	302 (6.2)	193 (6.7)	0.420	1634 (29.0)	463 (22.2)	**<0.0005**
Smoking	330 (48.9)	3268 (46.4)	0.213	1296 (49.2)	2473 (48.6)	0.596	1976 (49.5)	1868 (50.1)	0.562	2505 (51.8)	1449 (50.3)	0.229	3070 (54.5)	1165 (55.9)	0.274
Hypothyroidism	8 (1.2)	51 (0.8)	0.189	19 (0.7)	22 (0.4)	0.097	21 (0.5)	17 (0.5)	0.663	20 (0.4)	14 (0.5)	0.639	24 (0.4)	11 (0.5)	0.555
PVD	5 (0.7)	36 (0.5)	0.433	14 (0.5)	17 (0.3)	0.193	18 (0.5)	16 (0.4)	0.888	35 (0.7)	17 (0.6)	0.491	68 (1.2)	31 (1.5)	0.332
Year of cohort entry															
2000	165 (24.4)	2209 (31.4)	**<0.0005**	695 (26.4)	1629 (32.0)	**<0.0005**	989 (24.8)	878 (23.6)	0.219	637 (13.2)	304 (10.6)	**0.001**	334 (5.9)	102 (4.9)	0.08
2001	112 (16.6)	1867 (26.5)	**<0.0005**	612 (23.3)	1418 (27.9)	**<0.0005**	1022 (25.6)	914 (24.5)	0.284	1314 (27.1)	716 (24.9)	**0.028**	857 (15.2)	287 (13.8)	0.113
2002	113 (16.7)	1784 (25.3)	**<0.0005**	738 (28.0)	1237 (24.3)	**<0.0005**	1071 (26.8)	1028 (27.6)	0.445	1477 (30.5)	901 (31.3)	0.470	1618 (28.7)	578 (27.7)	0.391
2003	285 (42.2)	1185 (16.8)	**<0.0005**	587 (22.3)	804 (15.8)	**<0.0005**	912 (22.8)	906 (24.3)	0.125	1413 (29.2)	958 (33.3)	**<0.0005**	2826 (50.2)	1118 (53.6)	**0.007**
10 or more prior prescriptions	320 (47.4)	3232 (45.9)	0.446	1212 (46.1)	2188 (43.0)	0.011	1713 (42.9)	1614 (42.3)	0.704	2059 (42.5)	1260 (43.4)	0.290	2717 (48.2)	1004 (48.2)	0.961
Prior electrolyte ADRs	2 (0.3)	11 (0.2)	0.396	0 (0.0)	2 (0.04)	0.309	3 (0.08)	8 (0.2)	0.104	11 (0.2)	4 (0.1)	0.394	18 (0.3)	4 (0.2)	0.350
8 or more GP consultations	164 (24.3)	1814 (25.8)	0.409	658 (25.0)	1296 (25.5)	0.651	1098 (27.5)	1012 (27.2)	0.745	1433 (29.6)	807 (28.0)	0.141	1976 (35.1)	707 (33.9)	0.343

* Binary variables are presented as proportions (%), continuous variables are presented as mean (standard deviation); † Mann-Whitney test for non-parametric data

8.3.3.3 Matching on the propensity score

Standardized differences before and after matching using a caliper width of 0.001 are depicted in Figure 8.7 and presented in Table 8.5. Large standardized differences were observed in diabetes status, year of monitoring, SES score, sex, and antihypertensive treatment between patients with and without baseline testing prior to matching on the propensity score. After matching, there were no significant differences in the covariates between patients with and without baseline testing who were matched on the propensity score. The largest absolute standardized difference was 2.3%, suggesting good balance on the various covariates between patients with and without baseline testing in the matched sample.

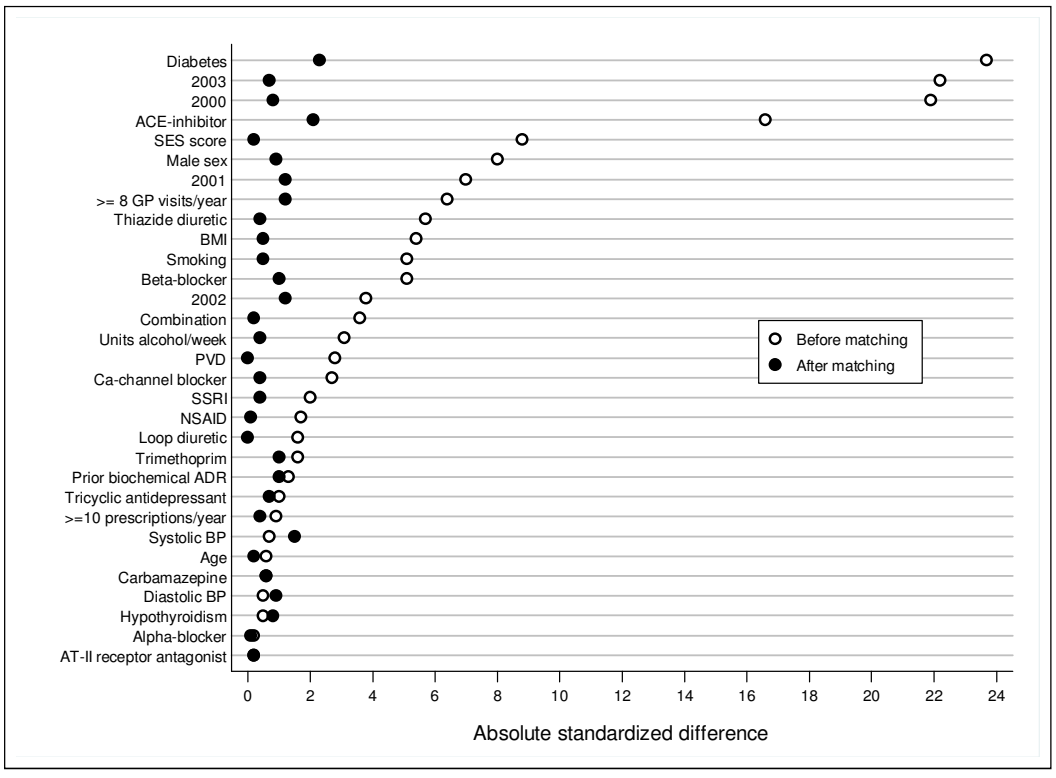


Figure 8.7 – Absolute standardized differences between patients with and without baseline testing before and after matching on the propensity score using a caliper width of 0.001

Table 8.5 – Standardized differences between patients with and without baseline testing before and after matching on the propensity score using a caliper width of 0.001

Characteristics	Before matching	After matching	% reduction in bias
Demographics			
Age	−0.6	−0.2	62.9
Male sex	8.0	0.9	89.4
SES score	8.8	0.2	98.2
BMI	5.4	0.5	90.2
Number of units of alcohol per week	3.1	0.4	88.1
Vital signs			
Systolic blood pressure	−0.7	−1.5	116.2
Diastolic blood pressure	−0.5	−0.9	66.3
Co-morbidities			
Diabetes	23.7	2.3	90.3
Smoking	5.1	0.5	90.3
Hypothyroidism	−0.5	0.8	70.8
PVD	2.8	0.0	100.0
Antihypertensive treatment			
ACE inhibitor	16.6	2.1	87.6
Alpha-blocker	0.2	0.1	47.4
AT-II receptor antagonist	0.2	−0.2	26.4
Beta-blocker	−5.1	−1.0	81.1
Ca-channel blocker	−2.7	−0.4	86.2
Combination	−3.6	−0.2	95.2
Loop diuretic	−1.6	0.0	100.0
Thiazide diuretic	−5.7	−0.4	92.5
Concomitant medications			
Carbamazepine	−0.6	0.6	-8.3
NSAIDs	1.7	0.1	94.1
SSRIs	−2.0	−0.4	78.6
Tricyclic antidepressants	−1.0	−0.7	33.4
Trimethoprim	1.6	1.0	36.8
Year of cohort entry			
2000	−21.9	−0.8	96.5
2001	−7.0	−1.2	82.7
2002	3.8	1.2	69.1
2003	22.2	0.7	97.0
10 or more prior prescriptions	0.9	−0.4	60.1
Prior electrolyte ADRs	1.3	1.0	18.5
8 or more GP consultations	6.4	1.2	81.4

A quantile-quantile plot (Figure 8.8) compares the propensity score distributions between patients with and without baseline testing before and after matching on the propensity score. Figure 8.9 presents the estimated propensity score distribution in patients with and without baseline testing in the unmatched and matched samples. A two-sample Kolmogorov–Smirnov test indicates that there is no significant difference in the distribution of propensity scores between patients with and without baseline testing in the matched sample (P=1.000).

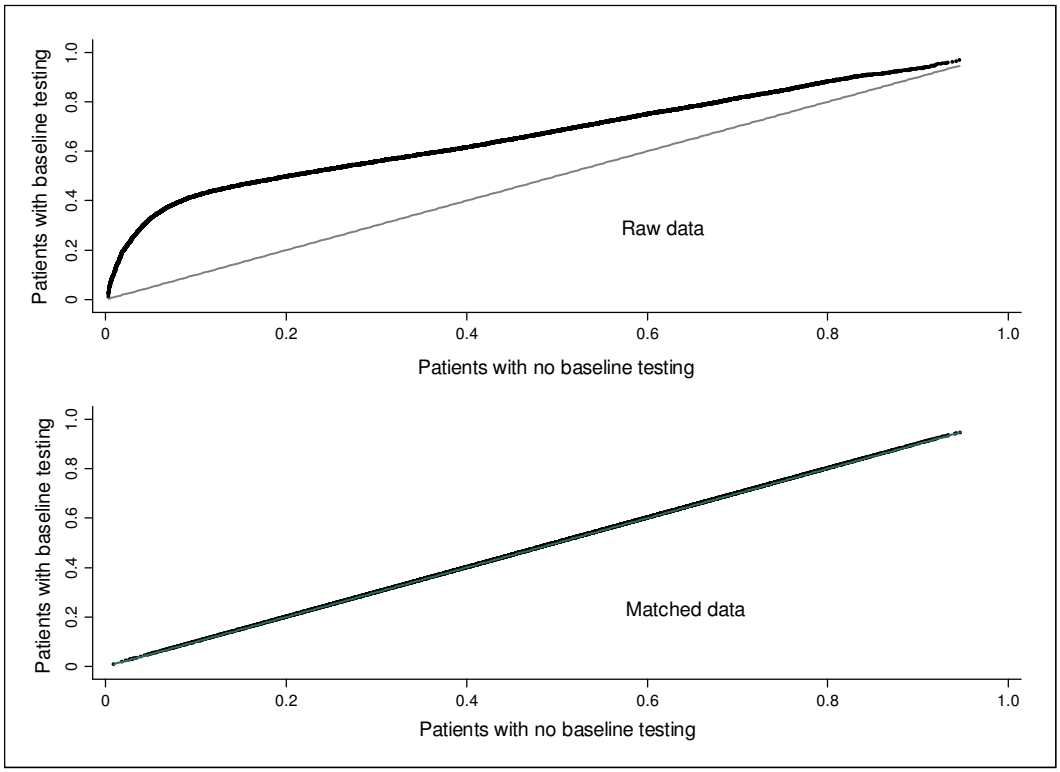


Figure 8.8 – Distribution of propensity score in patients with and without baseline testing before and after matching. The 45° line indicates identical distributions.

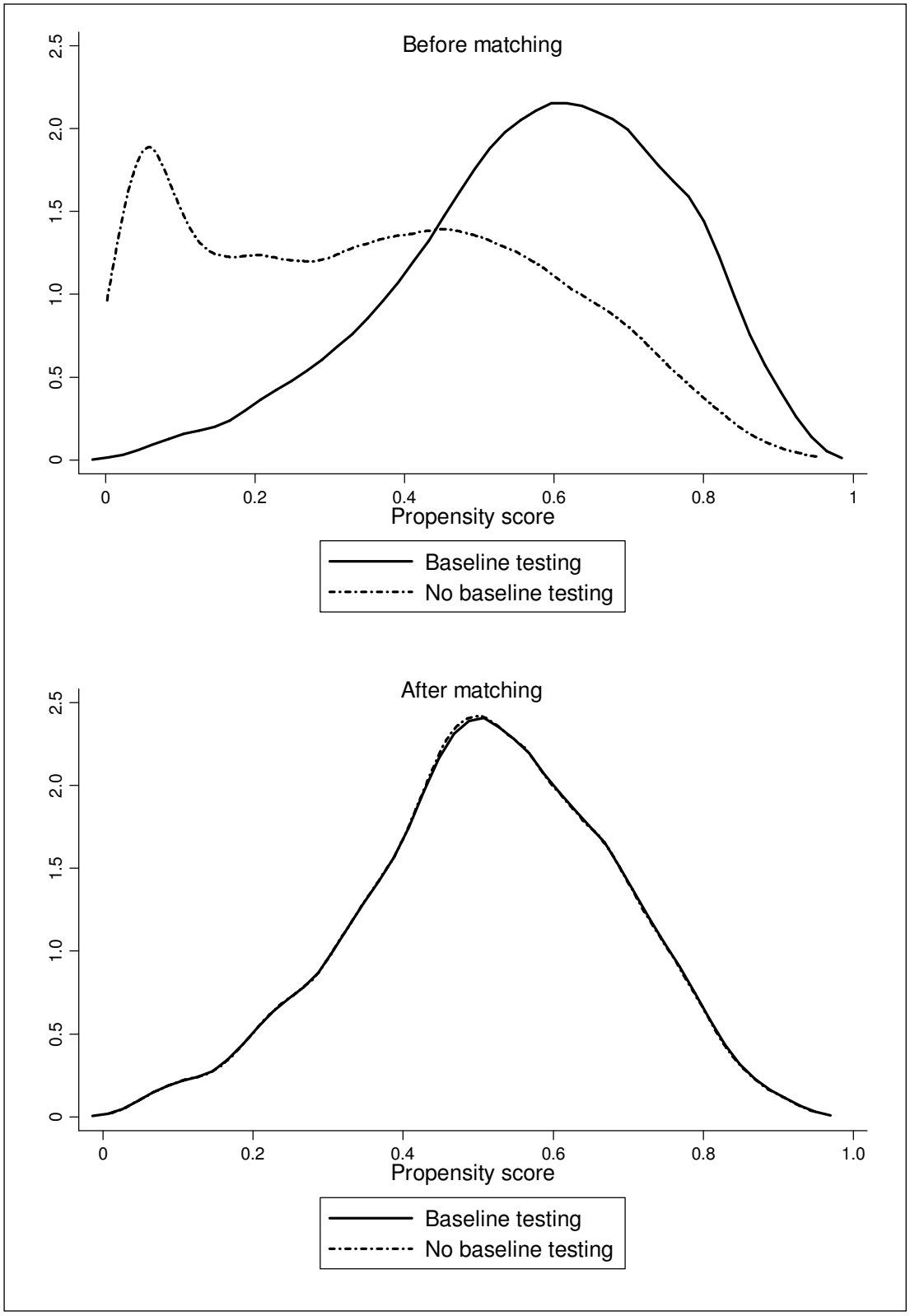


Figure 8.9 – Kernel density estimates of the propensity score in patients with and without baseline testing before and after matching on the propensity score

Patient characteristics in patients with and without baseline testing in the matched cohort are presented in Table 8.6. No significant differences between the two groups of patients were observed. In the matched sample, the largest absolute standardized difference was 2.3% (presented in Table 8.5), which also suggests good balance between patients with and without baseline testing.

Table 8.6 – Patient characteristics at baseline after propensity score matching

Characteristics	Any baseline testing (N=11795)	No baseline testing (N=11795)	P value
Demographics			
Age	59.8 (12.4)	59.8 (12.0)	0.862
Male sex	6489 (55.0)	6439 (54.6)	0.513
SES score	22.1 (16.6)	22.1 (16.6)	0.906
BMI	28.3 (5.1)	28.3 (5.1)	0.687
Number of units of alcohol per week	11.2 (14.4)	11.1 (14.5)	0.776†
Vital signs			
Systolic blood pressure	169.9 (19.1)	170.2 (20.1)	0.253
Diastolic blood pressure	97.5 (11.4)	97.6 (11.4)	0.512
Co-morbidities			
Diabetes	1001 (8.5)	922 (7.8)	0.060
Smoking	6023 (51.1)	5994 (50.8)	0.706
Hypothyroidism	67 (0.57)	60 (0.51)	0.533
PVD	76 (0.64)	76 (0.64)	1.000
Antihypertensive treatment			
ACE inhibitor	1833 (15.5)	1744 (14.8)	0.106
Alpha-blocker	126 (1.1)	125 (1.1)	0.949
AT-II receptor antagonist	235 (2.0)	238 (2.0)	0.889
Beta-blocker	2968 (25.2)	3017 (25.6)	0.463
Ca-channel blocker	832 (7.1)	843 (7.2)	0.780
Combination	109 (0.92)	111 (0.94)	0.892
Loop diuretic	43 (0.36)	43 (0.36)	1.000
Thiazide diuretic	5649 (47.9)	5674 (48.1)	0.745
Concomitant medications			
Carbamazepine	58 (0.49)	53 (0.45)	0.634
NSAIDs	1627 (13.8)	1623 (13.8)	0.940
SSRIs	390 (3.3)	399 (3.4)	0.744
Tricyclic antidepressants	442 (3.8)	457 (3.9)	0.610
Trimethoprim	195 (1.7)	180 (1.5)	0.435
Year of study entry			
2000	2238 (19.0)	2274 (19.3)	0.551
2001	2727 (23.1)	2787 (23.6)	0.356

Characteristics	Any baseline testing (N=11795)	No baseline testing (N=11795)	P value
2002	3324 (28.2)	3263 (27.7)	0.376
2003	3506 (29.7)	3471 (29.4)	0.618
10 or more prior prescriptions	5240 (44.4)	5262 (44.6)	0.773
Prior electrolyte ADRs	22 (0.19)	17 (0.14)	0.423
8 or more GP consultations	3359 (28.5)	3295 (27.9)	0.354

* Binary variables are presented as proportions (%), continuous variables are presented as means (standard deviation); † Mann-Whitney test for non-parametric data

Matching using the greedy algorithm and a caliper width of 0.001 created 11795 pairs of patients with and without baseline testing. Therefore, there were 5982 patients with baseline testing where no match could be found. These unmatched patients differed significantly from those patients with baseline testing who were matched successfully (Table 8.7). The unmatched group had a higher propensity score, and were more likely to be male, have a higher SES, a higher BMI, and consume more units of alcohol per week.

Table 8.7 – Baseline characteristics of matched versus unmatched patients with baseline testing

Characteristics	Matched patients (N=11795)	Unmatched patients (N=5982)	P value
Propensity score	0.512 (0.17)	0.704 (0.11)	<0.0005
Demographics			
Age	59.8 (12.0)	59.6 (11.7)	0.530
Male sex	6489 (55.0)	3599 (60.2)	<0.0005
SES score	22.1 (16.6)	23.8 (18.3)	<0.0005
BMI	28.3 (5.1)	28.7 (5.2)	<0.005
Number of units of alcohol per week	11.2 (14.4)	11.7 (14.1)	0.016†
Vital signs			
Systolic blood pressure	169.9 (19.1)	169.7 (18.8)	0.381
Diastolic blood pressure	97.5 (11.1)	97.3 (11.4)	0.213
Co-morbidities			
Diabetes	1001 (8.5)	1292 (21.6)	<0.0005
Smoking	6023 (51.1)	3154 (52.7)	0.036
Hypothyroidism	67 (0.57)	25 (0.42)	0.187
PVD	76 (0.64)	64 (1.1)	0.002

Characteristics	Matched patients (N=11795)	Unmatched patients (N=5982)	P value
Antihypertensive treatment			
ACE inhibitor	1833 (15.5)	1600 (26.7)	<0.0005
Alpha-blocker	126 (1.1)	57 (0.95)	0.471
AT-II receptor antagonist	235 (2.0)	128 (2.1)	0.511
Beta-blocker	2968 (25.2)	1212 (20.3)	<0.0005
Ca-channel blocker	832 (7.1)	344 (5.8)	0.001
Combination	109 (0.92)	26 (0.43)	<0.0005
Loop diuretic	43 (0.36)	19 (0.32)	0.616
Thiazide diuretic	5649 (47.9)	2596 (43.4)	<0.0005
Concomitant medications			
Carbamazepine	58 (0.49)	17 (0.28)	0.044
NSAIDs	1627 (13.8)	871 (14.6)	0.165
SSRIs	390 (3.3)	175 (2.9)	0.171
Tricyclic antidepressants	442 (3.8)	213 (3.6)	0.532
Trimethoprim	195 (1.7)	102 (1.7)	0.799
Year of study entry			
2000	2238 (19.0)	582 (9.7)	<0.0005
2001	2727 (23.1)	1190 (19.9)	<0.0005
2002	3324 (28.2)	1693 (28.3)	0.866
2003	3506 (29.7)	2517 (42.1)	<0.0005
10 or more prior prescriptions	5240 (44.4)	2781 (46.5)	0.009
Prior electrolyte ADRs	22 (0.19)	12 (0.20)	0.839
8 or more GP consultations	3359 (28.5)	1970 (32.9)	<0.0005

* Binary variables are presented as proportions (%), continuous variables are presented as mean (standard deviation); † Mann-Whitney test for non-parametric data

8.3.3.4 Effect of baseline testing

In the cohort of 38600 patients, only 76 patients (0.2%) died within six months of treatment. A greater proportion of patients discontinued their initial antihypertensive treatment (51%), was admitted to hospital (1.7%), or experienced an ADR (2.6%) (Table 8.8).

Table 8.8 – Adverse patient outcomes within six months of starting antihypertensive treatment

Adverse outcome, n (%)	Any baseline testing (N=17 777)	No baseline testing (N=20 823)	P value
Antihypertensive drug discontinuation	8 714 (49.0)	11 008(52.9)	<0.0005
Hospital admission	271 (1.5)	382 (1.8)	0.019
Death	32 (0.18)	44 (0.21)	0.489
Any ADR	464 (2.6)	536 (2.6)	0.824

A crude univariable analysis indicated that baseline testing was associated with a 17% decrease in the odds of hospital admission. No statistically significant association was demonstrated between baseline testing and death or baseline testing and any ADR (Table 8.9). Adjustment for potential confounding variables using multivariable logistic regression models suggested a larger decrease in the risk of hospital admission (OR 0.76, 95% CI 0.66–0.93) and a 28% decrease in the odds of the patient experiencing an ADR.

Methods using the estimated propensity score gave very similar results to those obtained from multivariable methods. Matching on the propensity score suggested that patients with baseline testing were significantly less likely to be admitted to hospital (OR 0.76, 95% CI 0.63–0.92) or suffer from an ADR (OR 0.70, 95% CI 0.60–0.81). When the regression model used the estimated propensity score as either a continuous variable or as a quintile, a significant decrease in the risk of hospital admission or developing an ADR within six months was demonstrated.

Table 8.9 – The relationship between baseline testing and adverse patient outcomes within six months

	Drug discontinuation		Hospital admission		Death		Any ADR	
	OR (95% CI)	**P value**	**OR (95% CI)**	**P value**	**OR (95% CI)**	**P value**	**OR (95% CI)**	**P value**
Propensity score (PS) methods								
PS matching (0.001 caliper)	0.90 (0.85–0.95)	<0.0005	0.76 (0.63–0.92)	0.005	0.88 (0.50–1.56)	0.662	0.70 (0.60–0.81)	<0.0005
Adjusted for continuous PS	0.90 (0.86–0.95)	<0.0005	0.78 (0.66–0.93)	0.006	0.73 (0.44–1.21)	0.221	0.72 (0.63–0.82)	<0.0005
Stratified by PS quintile	0.90 (0.86–0.94)	<0.0005	0.77 (0.65–0.92)	0.003	0.73 (0.45–1.21)	0.224	0.73 (0.64–0.84)	<0.0005
Within PS quintiles								
PS quintile 1	0.87 (0.74–1.02)	0.092	1.14 (0.59–2.19)	0.703	0.95 (0.12–7.36)	0.960	1.75 (0.95–3.25)	0.074
PS quintile 2	0.87 (0.79–0.96)	0.004	0.73 (0.51–1.04)	0.083	1.21 (0.39–3.70)	0.740	0.83 (0.60–1.14)	0.242
PS quintile 3	0.92 (0.84–1.01)	0.089	0.86 (0.62–1.19)	0.355	0.53 (0.22–1.27)	0.155	0.57 (0.44–0.75)	<0.0005
PS quintile 4	0.94 (0.85–1.03)	0.200	0.74 (0.53–1.04)	0.085	1.19 (0.36–3.95)	0.777	0.84 (0.65–1.08)	0.167
PS quintile 5	0.86 (0.77–0.95)	0.003	0.67 (0.45–0.99)	0.047	0.53 (0.20–1.39)	0.195	0.67 (0.51–0.87)	0.003
Traditional logistic regression (LR) methods								
Crude LR model	0.88 (0.85–0.92)	<0.0005	0.83 (0.71–0.97)	0.019	0.85 (0.54–1.34)	0.490	1.01 (0.89–1.15)	0.824
Multivariable LR model	0.90 (0.86–0.95)	<0.0005	0.76 (0.66–0.93)	0.006	0.72 (0.42–1.25)	0.247	0.72 (0.63–0.83)	<0.0005

Increasing the caliper widths increased the number of a patients with baseline testing that could be matched (Table 8.10). The corollary of this increase in sample size was that the decrease in the percentage bias achieved through matching was significantly diminished.

Table 8.10 – Standardized differences between patients with and without baseline testing using different caliper widths

Characteristics	Before matching	N=11 795		N=12 064		N=12 367		N=12 656		N=14 644		N=17 766	
		0.001 caliper	% reduction in bias	0.01 caliper	% reduction in bias	0.02 caliper	% reduction in bias	0.03 caliper	% reduction in bias	0.1 caliper	% reduction in bias	0.2SD logit PS caliper	% reduction in bias
Demographics													
Age	–0.6	–0.2	62.9	0.3	43.5	–0.7	11.0	0.4	32.9	–0.7	9.5	–1.0	62.7
Male sex	8.0	0.9	89.4	0.8	90.5	1.3	84.2	2.4	70.7	5.6	29.9	6.9	14.6
SES score	8.8	0.2	98.2	0.2	98.1	0.7	92.3	0.4	95.3	5.1	42.3	8.0	8.6
BMI	5.4	0.5	90.2	0.3	94.2	1.3	76.2	1.7	67.5	4.5	16.0	4.9	8.9
Number of units of alcohol per week	3.1	0.4	88.1	0.4	86.4	0.3	89.7	–0.2	92.3	1.6	49.2	2.3	28.1
Vital signs													
Systolic blood pressure	–0.7	–1.5	116.2	–0.7	3.2	–1.8	162.8	–1.3	86.7	–2.4	249.7	–1.6	126.1
Diastolic blood pressure	–0.5	–0.9	66.3	–0.5	2.7	–1.3	156.7	–1.2	131.8	–2.4	366.5	–0.9	76.5
Co-morbidities													
Diabetes	23.7	2.3	90.3	4.3	81.7	6.9	70.7	8.5	64.1	21.7	8.4	21.7	8.3
Smoking	5.1	0.5	90.3	0.1	98.7	0.7	85.3	0.3	93.4	2.3	55.2	4.2	17.4
Hypothyroidism	–0.5	0.8	70.8	–0.5	19.3	0.8	62.9	1.0	104.6	0.2	60.7	0.2	67.6
PVD	2.8	0.0	100.0	1.0	63.3	1.0	64.2	1.1	61.5	2.9	5.9	2.9	4.8
Antihypertensive treatment													
ACE inhibitor	16.6	2.1	87.6	2.8	82.9	5.8	65.2	5.6	66.4	13.8	16.9	14.4	12.8
Alpha-blocker	0.2	0.1	47.4	–0.2	2.9	–0.2	50.5	–0.6	292.2	–0.1	57.6	–0.1	65.1
AT-II receptor antagonist	0.2	–0.2	26.4	0.8	235.7	–0.9	274.3	–0.1	54.3	0.4	77.8	0.6	128.0
Beta-blocker	–5.1	–1.0	81.1	–1.2	76.6	–2.1	58.3	–2.3	54.3	–5.2	2.7	–5.0	1.7
Ca-channel blocker	–2.7	–0.4	86.2	–0.7	71.7	–1.2	55.7	–1.0	62.5	–2.9	9.3	–2.7	0.9
Combination	–3.6	–0.2	95.2	–0.3	92.9	–1.0	72.3	–1.4	61.6	–2.8	23.9	–3.3	8.3
Loop diuretic	–1.6	0.0	100.0	–0.1	91.5	0.1	91.7	0.3	83.9	–1.2	23.2	–1.7	9.3
Thiazide diuretic	–5.7	–0.4	92.5	–0.8	85.6	–1.3	76.5	–1.2	79.0	–3.6	36.9	–4.2	25.1
Concomitant medications													
Carbamazepine	–0.6	0.6	8.3	0.1	78.8	1.5	147.9	0.1	79.8	–0.4	30.2	–0.6	0.7
NSAIDs	1.7	0.1	94.1	–0.1	94.2	–0.4	73.3	0.5	72.6	0.7	58.5	1.6	6.2
SSRIs	–2.0	–0.4	78.6	–0.2	88.4	–0.1	93.2	–0.1	95.6	–0.8	59.8	–1.8	10.0
Tricyclic antidepressants	–1.0	–0.7	33.4	–0.7	34.9	–0.2	83.1	–0.5	50.4	–1.0	3.7	–1.2	23.8

Characteristics	Before matching	N=11 795		N=12 064		N=12 367		N=12 656		N=14 644		N=17 766	
		0.001 caliper	% reduction in bias	0.01 caliper	% reduction in bias	0.02 caliper	% reduction in bias	0.03 caliper	% reduction in bias	0.1 caliper	% reduction in bias	0.2SD logit PS caliper	% reduction in bias
Trimethoprim	1.6	1.0	36.8	0.7	58.8	−0.8	51.8	0.9	45.0	0.8	52.5	1.5	7.7
Year of cohort entry													
2000	−21.9	−0.8	96.5	−1.3	94.2	−1.9	91.3	−3.3	84.7	−13.4	38.7	−18.9	13.6
2001	−7.0	−1.2	82.7	−7.0	84.5	−1.8	73.7	−2.3	66.2	−4.7	33.1	−5.7	17.7
2002	3.8	1.2	69.1	0.6	85.1	0.0	100.0	−0.5	87.3	2.1	44.5	3.4	8.2
2003	22.2	0.7	97.0	1.6	92.9	3.4	85.3	5.6	74.6	14.2	36.1	18.7	15.7
10 or more prior prescriptions	0.9	−0.4	60.1	0.3	68.1	0.4	60.2	1.1	14.9	2.7	189.2	1.9	103.5
Prior electrolyte ADRs	1.3	1.0	18.5	0.6	52.2	0.0	100.0	0.2	84.8	1.3	5.1	1.2	2.6
8 or more GP consultations	6.4	1.2	81.4	1.5	76.9	2.2	66.1	2.8	56.6	5.5	15.1	5.9	8.3

SD = standard deviation

Table 8.11 – The relationship between baseline testing and adverse patient outcomes within six months using propensity score matching with different caliper widths

Propensity score caliper width	Drug discontinuation		Hospital admission		Death		Any ADR	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
0.001 caliper	0.90 (0.85–0.95)	<0.0005	0.76 (0.63–0.92)	0.005	0.88 (0.50–1.56)	0.662	0.70 (0.60–0.81)	<0.0005
0.01 caliper	0.90 (0.85–0.94)	<0.0005	0.76 (0.63–0.92)	0.004	0.67 (0.37–1.21)	0.183	0.72 (0.62–0.83)	<0.0005
0.02 caliper	0.89 (0.84–0.93)	<0.0005	0.73 (0.60–0.88)	0.001	0.68 (0.38–1.21)	0.192	0.69 (0.60–0.80)	<0.0005
0.03 caliper	0.89 (0.85–0.94)	<0.0005	0.76 (0.63–0.91)	0.003	0.70 (0.40–1.22)	0.210	0.77 (0.67–0.89)	<0.0005
0.1 caliper	0.87 (0.83–0.92)	<0.0005	0.72 (0.61–0.86)	<0.0005	0.82 (0.49–1.36)	0.439	0.78 (0.68–0.90)	<0.0005
0.2SD logit PS caliper	0.88 (0.84–0.92)	<0.0005	0.76 (0.65–0.89)	0.001	0.80 (0.50–1.27)	0.347	0.89 (0.78–1.00)	0.067

8.4 Discussion

Patients who received baseline testing of serum electrolytes or renal function prior to the initiation of antihypertensive therapy were younger, more likely to be male, and more likely to have diabetes. Both traditional multivariable regression methods and propensity score methods demonstrated a decreased risk of hospital admission and any ADR in patients who had a record of any baseline testing. Similarly, both methods indicated that there was no association between baseline testing and death.

I chose to use propensity score methods as they may generate estimates that are more precise and less biased than traditional multivariable methods when the number of events per confounder is low (Cepeda *et al.*, 2003). These methods can control for measured confounding in observational research and allow for the assessment whether balance has been achieved between patients with and without the treatment of interest. Propensity score methods may also provide additional benefit by restricting the analysis to groups that overlap on potential confounders, making the groups more comparable (Hill *et al.*, 2004).

The imbalance in the observed covariates was reduced between patients with and without baseline testing using both stratification and matching on the propensity score. Stratification using propensity score quintiles, which used the entire sample, did not yield very different results from one-to-one matching on the propensity score. Matching allowed for the reduction in the imbalance in patient covariates between patients, although this process decreased the sample size from which an estimate of the effect of baseline testing on adverse patient outcomes could be measured. I initially used a caliper of 0.001 for matching, which is narrower than the caliper

widths recommended elsewhere (Austin, 2009c). The tighter caliper allowed for improved balance between potential confounding variables, such as diabetes mellitus, but further decreased the sample size. As a sensitivity analysis, various caliper widths were used to examine the effect of the using wider widths on the results. Wider caliper widths increased the sample size, decreased the balance in the patient covariates between the two groups, but did not cause a very large difference in the results.

The development of the propensity score was done using an iterative method. Variables were selected *a priori* if they were believed to be associated with baseline testing but also if they were thought to be associated with the adverse outcome. The inclusion of variables that are related to the outcome, irrespective of their relationship between the exposure, has been previously suggested (Rubin, 1997). Brookhart and colleagues (2006) demonstrated that the inclusion of variables that are related to the exposure but not to the outcome increases the variance of the estimated exposure effect.

The propensity score was evaluated using graphical methods and standardized differences to assess the ability of the score to achieve balance on potential confounders. The graphical methods indicated that the estimated propensity score achieved good balance on measured confounders, although there was evidence of some residual imbalance in the extreme quintiles of the propensity score. Although some authors have previously used the Hosmer–Lemeshow goodness-of-fit (GoF) test or the area under the receiver operator characteristic (ROC) curve (the *c*-statistic) as a way to measure the adequacy of the propensity score (Weitzen *et al.*, 2004), these

tests were not used for several reasons. The GoF test and the *c*-statistic do not provide any information on detecting important missing confounders in a propensity score model (Weitzen *et al.*, 2005). Moreover, Austin (2009a) demonstrated using a simulated data set that the c-statistic could not differentiate between a propensity score model that had been misspecified and one that had been correctly specified.

Propensity score methods have been shown to be robust and perform well when the outcome of interest is rare. However, they may slightly decrease the strength of the association between treatment and outcome. Although there has been a significant increase in the use of propensity score methods over the past several years, research has demonstrated that they continue to generate similar results to those obtained from traditional regression models (Shah *et al.*, 2005; Stürmer *et al.*, 2006a), as was observed in my analysis. In this case, this may have been due to the relatively large sample size and a sufficient number of adverse patient outcomes that were not rare.

8.4.1 Limitations

Matching based on the propensity score created excellent balance between patients with and without baseline testing on potentially confounding variables. However, matching significantly reduced the sample size as 5982 patients with baseline testing were lost because no match could be found. These patients were significantly different from patients for whom a match could be found, which may have introduced some bias into the results. The use of propensity scores, as well as traditional multivariable modelling, is also limited by the fact that they cannot control for confounding due to unmeasured variables.

Finally, the analysis was restricted to patients with complete data on patient covariates. Because propensity scores are estimated using a wide range of covariates, a number of patients may have been excluded from the analysis as a significant proportion of patients often have missing values for at least one covariate. Patients with missing data may have therefore differed from those patients with complete data, which may have introduced bias into the results. Two methods have been suggested for implementing propensity scores methods with multiple imputed data, although few studies have undertaken this work (Hill, 2004). In the first method, propensity scores are calculated for each imputed data set and the propensity scores are then combined across the imputed data sets. A matched control group is then selected allowing for the calculation of the effect estimate. In the second method, the propensity scores are again calculated for each imputed data set but in this method the matched control group is selected within the imputed data set and the effect estimate is calculated. The effect estimates are then combined across all of the imputed data sets.

8.4.2 Potential further statistical analysis using propensity score methods

I used propensity score methods to control for confounding by indication in my analysis of baseline testing and adverse patient outcomes. However little work has been carried out to investigate how propensity score methods should be used when there is potential for immortal time bias (which has been described in detail in section 4.3). This would have been required if I had modelled the propensity of receiving follow-up monitoring.

Several approaches have been suggested for using propensity score methods with time-dependent exposures. A Cox proportional hazards model where survival time is regressed on the propensity score and treatment status has been used, where treatment is modelled as a time-varying covariate (Tleyjeh *et al.*, 2010). A second approach has been to use the follow-up time as a matching variable, which ensures that treated patients are matched with untreated patients who survived at least as long as the time to the event in a matched, treated patient (Kumar *et al.*, 2010; Lalani *et al.*, 2010; Tleyjeh *et al.*, 2010). A third approach that has been used to deal with both selection bias and immortal time bias is to restrict the analysis to patients who experienced the time-dependent event at baseline or time of study entry (e.g. restricting an analysis that examines the relationship between treatment and an outcome to patients who had the treatment on the day of hospital admission) (Lindenauer *et al.*, 2010).

8.4.3 Conclusion

I examined the relationship between baseline testing and adverse patient outcomes using different propensity score methods and a traditional multivariable regression technique. Both techniques found that patients with baseline testing had a significantly decreased risk of hospital admission or the development of an ADR within six months of starting antihypertensive treatment.

Chapter 9

CONCLUSIONS AND FUTURE WORK

This chapter summarises the important study findings and describes the important strengths and limitations, and generalizability of the results. I also describe future research opportunities that need to be addressed.

9.1 Background to the research

The monitoring of drug treatment has been suggested as a way to improve therapy decisions, allow for better titration of treatment, improve adherence to therapy, and identify potential adverse reactions to treatment. Monitoring for ADRs that may not necessarily manifest symptoms is of particular importance.

Antihypertensive therapy is commonly used in general practice but is known to cause a range of ADRs, particularly electrolyte disturbances. Published guidelines recommend that patients treated with antihypertensive therapy should have tests of their renal function and electrolyte concentration in order to identify biochemical ADRs. It is not well known to what extent these guidelines are followed in primary care.

As a consequence, a systematic review was undertaken in order to identify studies examining the nature of monitoring for ADRs. I identified several large, well-planned studies that examined the prevalence of monitoring in patients treated with antihypertensive drugs and identified patient factors associated with greater rates of monitoring. The estimated rate of baseline biochemical testing in primary care differed markedly, ranging across studies from 17% to 81%. The proportion of patients with follow-up monitoring also differed, from 20% to 79%. These wide ranges may reflect differences in the rates of monitoring, differences in the methods and definitions of monitoring used by the studies, or random error. In addition, very few studies examined both baseline testing and follow-up monitoring, and only one previous study examined relationships between monitoring and adverse patient outcomes.

9.2 Summary of the findings

I used data from the GPRD—the world's largest database from primary care— to provide specific insight into the nature of monitoring for biochemical ADRs in a cohort of over 74000 patients newly treated with antihypertensive therapy. The main findings from this thesis are summarized in relation to the original aims of the study.

1. *To observe and clarify the patterns of monitoring for ADRs in patients treated with antihypertensive drugs;*

The study demonstrated that monitoring for ADRs in patients treated with antihypertensive drugs in primary care often did not occur. Baseline biochemical testing, which was defined as any test in the six months prior to the initiation of

antihypertensive therapy, was recorded in 42% of the cohort. Only half of the patients had any laboratory test within one year of starting antihypertensive treatment, and 36% had any laboratory test within six months. Few patients—only 1 in 5—had more than one follow-up laboratory test. The scarcity of tests must limit the ability of GPs to monitor changes in renal function or electrolyte concentration during antihypertensive treatment. Less than one quarter of patients had both baseline testing and follow-up monitoring. This too must limit the ability of clinicians to assess intra-individual changes in serum concentrations after the initiation of treatment.

Similar low levels of monitoring where observed when I examined a sub-group of thiazide diuretic-treated patients to provide specific insight into electrolyte monitoring and the development of hyponatraemia in a cohort of patients treated with the most common type of antihypertensive therapy. Only 37% had electrolyte testing prior to the initiation of thiazide diuretic treatment and 46% had any electrolyte monitoring during one year of treatment.

In patients with records of any biochemical monitoring, the time between laboratory tests varied significantly and very few patients demonstrated 'regular' monitoring, where the time intervals between tests were approximately equal. Follow-up monitoring did not increase after an abnormal test result, dose change, or co-prescription of other drugs that could lead to potentially harmful drug-drug interactions – events where guideline have recommended increased levels of monitoring. Indeed, a decrease in the density of follow-up laboratory tests (the number of tests over time) was observed.

2. *To determine relationships between the outcomes for patients, the patterns of monitoring for ADRs, and other factors;*

Patient factors such as treatment with an ACE-inhibitor, increasing age, the year of initiation of antihypertensive treatment, and diabetes mellitus were shown to be strongly associated with follow-up biochemical monitoring. These same factors were shown to be associated with follow-up monitoring when allowances were made to deal with potential bias introduced by missing data on important patient covariates. The increase in monitoring with patient age and disease may suggest that GPs tend to adhere to monitoring guidelines for certain patients to whom they may attribute a greater risk of harm.

I modelled the relationship between follow-up monitoring and three adverse patient outcomes: antihypertensive drug discontinuation, hospital admission, and death. Follow-up monitoring was treated as a time-varying covariate in order to control for immortal time bias. Patients who had any follow-up monitoring were slightly but significantly more likely to experience all three outcomes. Death was rare and only a small proportion of patients died within one year of starting antihypertensive treatment. Patients with follow-up monitoring were statistically more likely to die, although the small numbers of death implies a greater uncertainly in the results. Less than 4% of patients were hospitalized within one year of treatment. Patients with follow-up monitoring may have been more likely to be admitted to hospital for several reasons: the doctors ordered some follow-up monitoring when the patient presented with an illness that would later progress to hospital admission, patients were

admitted to hospital as a consequence of the results of a test, or patients had a test as prelude to admission to hospital for another reason.

Discontinuation of antihypertensive treatment was common, with over half of the patients stopping treatment within one year. Indeed almost one in five patients had a discontinuation date of 28 days, suggesting that a large proportion of patients discontinued treatment after one course of therapy. Monitoring was associated with an increased risk of antihypertensive discontinuation, which may have been associated with increased contact with the GP that allowed for the further discussion on the nature of the treatment, lack of drug efficacy, or a serum concentration value that warranted discontinuation of treatment.

I also determined the rate of biochemical ADRs in the year following the initiation of treatment. Fewer than 1 in 20 patients had a record of a biochemical ADR. The majority of the ADRs were identified only through the use of a certain serum concentration value outside of a given range and not by a clinical code.

I undertook a sub-group analysis of thiazide diuretic-treated patients—the most commonly used antihypertensive therapy—to examine the development of hyponatraemia. I demonstrated that one in ten patients with serum sodium monitoring had evidence of hyponatraemia during thiazide diuretic treatment. The majority of cases of hyponatraemia were mild and most often normalized on the subsequent serum sodium concentration test. A relationship between hyponatraemia and hypokalaemia was demonstrated as patients with evidence of severe hyponatraemia were significantly more likely to have a serum potassium concentration test indicating

hypokalaemia. Patients with evidence of hyponatraemia at baseline remained hyponatraemic upon subsequent monitoring after initiation of thiazide diuretic treatment, which should perhaps encourage more baseline testing. An increased risk of hyponatraemia was identified in older patients, those with a low body mass index, female patients, and those taking concomitant carbamazepine. The increased risk in women may be due to several factors including smaller BMI, increasing age, a lower dietary intake of sodium, an over-representation of women in the thiazide diuretic-treated cohort, or because women are known to have more ADRs than men.

3. *To compare the observed patterns of monitoring of patients after the initiation of antihypertensive drugs for ADRs against published recommendations for monitoring;*

It was difficult to make comparisons between the patterns of monitoring in the GPRD and published recommendations for monitoring because of the varied nature of the recommendations. The majority of published guidelines identified recommended baseline testing prior to starting antihypertensive therapy and at least some follow-up monitoring after the start of treatment. However, some guidelines specifically stated that follow-up monitoring should be undertaken one week and one month after starting treatment, and at intervals following any dose changes. Even when the less frequent of recommendations was used as reference, the patterns of testing and monitoring by UK GPs in this cohort of patients did not follow these guidelines. Indeed, only half of patients and 16% of patients had any monitoring in the year and one month after initiation of antihypertensive therapy. A change in dose was not

associated with increased monitoring; indeed, the density of follow-up laboratory tests was significantly smaller after an increase in the dose of antihypertensive therapy.

4. Additional findings supplementary to the original aims;

It became clear after the original aims of the thesis were envisioned that further work using more advanced statistical and pharmacoepidemiological techniques was desirable. The proportion of missing data was a significant concern, because of the potential for bias if the data were not missing completely at random. I demonstrated that several patient factors such as age and sex were associated with the missingness of different patient covariates, which suggested that an analysis restricted to patients with complete data may have been biased. When the primary analysis investigating the patient factors associated with follow-up monitoring was carried out using multiple imputed data, similar results were obtained with those from the complete case analysis. However, the results using the multiple imputed data had narrower confidence intervals and more precision.

The issue of confounding and how one can control for confounding in observational research was another concern. I chose to use propensity score methods to control for confounding in the analysis of the potential relationship between baseline biochemical testing and adverse patient outcomes within six months of treatment. Propensity score methods were first fully described in 1983 but have been used more widely in recent years. Their use in pharmacoepidemiological research has grown substantially, mainly because they have been shown to perform particularly well when there are few events per potential confounding variable.

Patients with evidence of serum electrolyte or creatinine tests prior to antihypertensive treatment (baseline testing) were at lower risk of discontinuation of antihypertensive therapy or a biochemical adverse drug reaction within six months of starting antihypertensive treatment, than patients who had no evidence of baseline biochemical tests. Propensity score methods generated more precise although still similar results compared with those from traditional multivariable methods.

5. Examining the different findings from chapters 5 and 8 investigating the association between adverse patient outcomes and baseline biochemical testing and follow-up biochemical monitoring.

I differentiated between biochemical tests recorded prior to the start of treatment (baseline testing) and tests started after the patient has been treated with antihypertensive therapy (biochemical monitoring). Tests prior to the initiation of therapy can discover secondary causes of hypertension and coincidental baseline abnormalities; and act as a benchmark against which to assess change, while tests taken during treatment can identify changes in serum concentration resulting from adverse reactions to the therapy. I investigated the potential association between adverse patient outcomes and follow-up biochemical monitoring (Chapter 5) and again between adverse patient outcomes and baseline testing (Chapter 8).

The risk of adverse patient outcomes was significantly lower in patients with any record of biochemical tests in the six months prior to the initiation of antihypertensive therapy than in those with none. These results are the converse of those presented in Chapter 5 where follow-up monitoring after the start of treatment was associated with

an increased risk of adverse outcomes. This relationship between baseline testing and decreased risk of adverse outcomes may, in fact, be an artefact of the way the data were modelled and not represent a true association. Indeed, I demonstrated that patients with baseline biochemical testing were significantly less likely to have any follow-up monitoring in the six months after the initiation of antihypertensive therapy (adjusted OR 0.80 (95% CI 0.77–0.83)). This suggests that patients with baseline testing may have been seen less frequently by their GP, thus reducing the number of opportunities to undertake reactive monitoring (i.e. where the patient presented to the practice with an illness requiring monitoring that would lead to the adverse outcome). The lack of follow-up monitoring in patients with baseline testing may have led to a decrease in contact with the GP and concomitant decrease in the adverse patient outcomes that I examined, therefore causing a potentially inverse association between baseline testing and the outcomes.

Intuitively, monitoring should be beneficial. However, the retrospective analysis of electronic health records does not allow the true relationship between monitoring and adverse patient outcomes to be unravelled; that will require prospective studies, as I discuss in section 9.4.1.

9.3 Strengths and limitations of the thesis

I have presented the strength and weaknesses of the various analyses in the previous chapters. Here, I describe the overall strengths and limitations of the thesis. This study is strengthened by its large sample size and access to a range of baseline patient covariates. Earlier work in the UK has been carried out in significantly smaller populations. In addition, the majority of previous research on biochemical monitoring

using large scale studies has been limited primarily to HMOs in the United States, where monitoring is often driven by targets, and therefore the results of such studies may not be relevant to the UK.

Missing data are a common problem in studies of primary care databases and I demonstrated that there were significant associations between patient factors and the missingness of the data. This is one of the few studies using the GPRD to undertake multiple imputation to control for the potential bias that can occur when complete case analyses are carried out.

Finally, this is only the second study to investigate the relationship between follow-up monitoring and adverse patient outcomes. The data were analysed in such a way as to control for immortal time bias and I demonstrated how failure to control for this bias created an artificial protective effect.

Certainly, the main limitation of this thesis lies in the assumption that a record of a laboratory test was evidence of monitoring for an ADR. Because the data were obtained from a retrospective analysis of the GPRD, there was no way of knowing what the impetus was for ordering the laboratory test or the context in which monitoring occurred. I demonstrated an increased risk of adverse patient outcomes in patients with monitoring, which I have interpreted to mean that monitoring was actually *reactive* instead of being *planned* (i.e. carried out in accordance with published guidelines). However, this conclusion is only an assumption as the results are based on a retrospective analysis of records, and it is impossible to precisely determine whether the monitoring itself was truly planned or reactive. In addition, I

cannot be sure whether the laboratory test was ordered specifically because the patient was being treated with an antihypertensive drug. Finally, it was impossible to differentiate between GPs failing to undertake biochemical monitoring and patients failing to attend for laboratory testing.

The results may also be limited by the time frame of the analysis. I demonstrated that follow-up monitoring improved over time, with 63% of patients who started antihypertensive treatment in 2003 having any monitoring compared with only 38% of patients started treatment in 2000. This is most likely due to an increase in the number of practices becoming linked electronically to laboratories. This electronic link allows test results to be incorporated automatically into the GP practice's electronic records. Without this direct link to the laboratory, it would have been up to the practice to record paper-based laboratory results in their electronic records, which may have caused an under-reporting of monitoring. This increase in the level of monitoring may also reflect, to a lesser extent, the general trend in United Kingdom general practice towards improving quality standards and recording. Because the data were obtained from an earlier time period, it is impossible to determine whether the data presented in this thesis reflect current practice in biochemical monitoring. However, the differences in monitoring between drug classes and among patients are likely to persist.

Finally, I did not attempt to assess any potential effect of clustering based on the GP practice. The mean number of patients each practice contributed to the cohort was 185, but the results were skewed with five practices contributing fewer than ten patients and six practices contributing more than 500 patients. Differences between

GP practices can exist due to different standards, staffing levels and general practitioner characteristics, which can impact upon the extent of laboratory monitoring that is undertaken. Indeed, previous research has demonstrated significant differences in the demographic and professional characteristics of GPs in the number and type of laboratory tests ordered (Vinker *et al.*, 2007). Matching patients on the GP practice would have, in theory, decreased this risk of confounding. However, significant information would have been lost if no match was found.

9.4 Should we really monitor for ADRs to antihypertensive therapy?

Monitoring tests have become a major element of primary care and intuitively, monitoring should be beneficial (Glasziou, 2007). Indeed, doctors have described monitoring as a critical component of their practice, although they also view it as a time-consuming process (Goldman *et al.*, 2010). It is therefore necessary to determine whether the benefits of monitoring outweigh any harms such as inconvenience and cost, and the potential impact that false positive or false negative results may have on treatment (Glasziou *et al.*, 2005). Certainly, false positive results can lead to the potentially harmful 'ping-pong' effect [referred to as 'hunting' in control theory], where changes in treatment in response to a test are too large and can increase with the variation in the monitored variable (Glasziou & Aronson, 2008). This effect may not be as important in the monitoring for adverse reactions to antihypertensive treatment as it would be in the therapeutic monitoring of a patient's INR during warfarin treatment, but still is a potential cause for concern in a monitoring strategy. One also doesn't want to measure too often, as apparent changes in test results may only be the result of short-term biological variation and analytical variation.

I used a retrospective analysis of observational data to investigate the nature, frequency, and responsiveness of biochemical monitoring in clinical practice. I demonstrated that only half of patients had any follow-up monitoring within one year of antihypertensive treatment. Few patients had more than one monitoring test, which may suggest that GPs used a 'hit and run' type of monitoring, where the patient was tested once—usually to assess the initial response to therapy—and not again (Glasziou & Aronson, 2008). I also demonstrated an increased risk of adverse patient outcomes in patients who were monitored, which may simply be a reflection of a more reactive approach to monitoring whereby GPs order tests when a patient presents with a condition that requires hospital admission or discontinuation of treatment.

I was also able to demonstrate that 9% of serum concentration tests recorded within one year of antihypertensive treatment were outside of the standardized reference ranges. These tests identified patients that may have been experiencing ADRs that may not have manifested any symptoms and prevented drug-induced harm. However, the corollary of this is that a large proportion of tests (91%) were normal. In addition, even though it was not possible to determine the false positive rate, a large proportion of the patients with an abnormal serum concentration had a subsequent test that returned to normal. Specifically, almost three quarters of patients with an abnormal potassium test had a subsequent normal potassium test, which would suggest that either the tests of serum concentration generate a number of false positive results or that there is considerable within-patient variance in the serum concentrations.

This retrospective analysis of observational data was able to characterize and describe the nature of biochemical monitoring for ADRs, and provide specific insight into monitoring in primary care in the UK. However, it cannot be used to answer the question of whether doctors should monitor at all for ADRs. Furthermore, this type of analysis can neither be used to estimate the benefit of monitoring nor to create a definitive evidence-based monitoring strategy for ADRs in patients treated with antihypertensive therapy. Instead, randomized controlled trials or prospective studies of monitoring are required to determine whether monitoring for adverse reactions to antihypertensive therapy itself is a rational and useful exercise.

9.4.1 What types of studies can be used to determine whether we should monitor at all for ADRs?

In general, randomized controlled trials or prospective studies of monitoring are limited, although there have been some randomized trials examining the benefits of therapeutic drug monitoring. For example, Jannuzzi and colleagues (2000) demonstrated in a study of patients with epilepsy that the mean serum concentrations of antiepileptic drugs outside the target range was significantly lower in the monitored group compared with the control group (8% vs. 25%), but the proportion remaining free of seizures did not change (38% vs. 41%). Another small randomized trial investigated the monitoring of plasma concentrations of HIV protease inhibitors and demonstrated that monitoring improved the number of patients with a low viral load (Burger *et al.*, 2003). There have also been modelling studies that have investigated the benefits of monitoring the effects of treatment such as monitoring blood pressure during ACE inhibitors treatment (Bell *et al.*, 2010) or monitoring bone mineral density with bisphosphonate treatment (Bell *et al.*, 2009).

No prospective observational studies or randomized trials of monitoring for biochemical ADRs in patients treated with antihypertensive therapy have been undertaken. The majority of studies examining monitoring for ADRs, like the one presented in this thesis, have been retrospective in design. Certainly data on adverse reactions to treatment from clinical trials exist and these provide the evidence on the epidemiology of ADRs and how they are manifest in terms of dose, time, and various susceptibility factors. However, these types of studies were primarily designed to evaluate the efficacy and safety—although perhaps to a lesser extent—of antihypertensive therapy. Monitoring was not the focus of these clinical trials and the data on ADRs are often applied to monitoring guidelines, without consideration to the context of the monitoring.

There is limited knowledge of the signal-to-noise ratio of the laboratory tests for renal function and electrolyte concentration. Ideally, one would want a monitoring test with a high signal-to-noise ratio, where the test is able to differentiate changes due to an ADR from background measurement variability caused by short-term biological fluctuations and technical measurement error (Mant, 2008). The average long-term change in the serum concentration (the signal) and the short-term within-person variation (the noise) need to be understood in order to determine whether the monitoring tests are appropriate and to better inform the development of a monitoring strategy. Certainly, there was some indication in the analysis of the GPRD data that the serum concentration tests may have a low signal-to-noise ratio as a large number of the abnormal serum potassium tests returned to normal. This may be due to problems with validity of the tests or due to other considerations such as problems with storage and how the tests were obtained. For example, a serum potassium

concentration can be artificially increased if fist clenching or pumping is used to obtain the blood sample or if the blood sample is stored for a longer period of time. It is therefore necessary to determine the number of monitoring tests that are true positives and to then model the proportion of tests that need to be true positives before monitoring can be regarded as worthwhile.

It is also necessary to determine the appropriate timing and frequency of monitoring tests. Deciding on the test frequency, particularly at the initiation of treatment, requires an understanding of the dose of treatment and speed at which changes in the serum concentrations may develop (Coleman *et al.*, 2006).

I propose two types of studies to develop the primary evidence base for the monitoring of adverse reactions to antihypertensive therapy and to better address the question of whether doctors should monitor at all for ADRs. The first study would be an experimental trial where biochemical monitoring of patients occurred as frequently as practically and ethically feasible (e.g. once a week or once a fortnight) in order to identify potential ADRs. Observations could then be deleted in order to compare the specificity or specificity of the different monitoring regimens (e.g. monitoring every week versus monitoring every four weeks). For example, if 90% of all pre-specified ADRs were identified by monitoring every week, compared with 87% of ADRs that were identified by monitoring every month, this would suggest that the benefit from monitoring every week compared with every month would be slight.

The second study would be a modelling exercise using trial data and analogous to the study by Glasziou and colleagues (2008), who investigated serum cholesterol

concentration monitoring in patients treated with pravastatin. Data would be obtained from a randomized trial where antihypertensive therapy was compared with placebo and where tests of renal function or electrolyte concentrations were taken before the initiation of treatment and at regular intervals subsequently. For example, the HOPE study—a large RCT of ramipril—measured both creatinine and potassium pre-randomization, at one month after the start of treatment, and at yearly intervals (Yusuf *et al.*, 2000). The analysis would need to estimate several parameters. First, the variation in the change in serum concentration from baseline to initial monitoring (e.g. one month or three months) would be estimated. Second, the extent to which the long-term change varies within and among patients can be estimated. Finally, the signal-to-noise ratio for the serum concentration test would be calculated. The average long-term change in the serum concentration between treated and untreated patients (the signal) would be compared with the short-term within-person variation (the noise), which can be calculated using the pre-treatment measures of the serum concentrations.

This type of modelling exercise would allow for a comparison between treated and untreated patients in the variation in the initial response, long term changes in the serum concentrations and perhaps most importantly, the ratio of the long term changes to the within-person variation. These parameters would generate a better estimate of the frequency of monitoring required to identify true changes in serum concentrations and the potential risk of biochemical ADRs.

9.5 What are the barriers to monitoring?

9.5.1 Lack of consensus in monitoring guidelines

One of the barriers to monitoring is the lack of consensus between expert and national hypertension societies in the guidelines for monitoring, which is likely due to the poor evidence base for monitoring. Some published guidelines are vague and describe monitoring as a 'routine investigation' (Williams *et al.*, 2004), while only a small number of guidelines provide detailed recommendations for biochemical testing prior to the initiation of therapy, specific details for the frequency of follow-up monitoring, and actions to be taken should the laboratory tests be outside a certain range of values (French Haute Autorité de Santé, 2005; Smellie *et al.*, 2007). Indeed, monitoring guidelines tend to be subjective and based solely on expert opinion (McAlister *et al.*, 2007). When Eccles and colleagues (1998) developed a monitoring guidelines for ACE inhibitors used in the treatment of heart failure they stated that they 'could find no basis for recommending one monitoring interval over another in long term treatment, and felt that monitoring at least once a year was appropriate'. This uncertainty can lead doctors to making decisions on monitoring based on their clinical experience rather than evidence, which may in turn create opportunities for adverse events (Goldman *et al.*, 2010).

9.5.2 Absence of alerts or reminders to monitor

Another potential barrier to monitoring is the lack of well-designed and implementable electronic alerts to remind the practitioner of the need to undertake monitoring both pre- and post-initiation of therapy, as well as alerts to indicate that a

patient had not attended for monitoring. Considerable recent research has been carried out to examine the impact of electronic alerts and reminders to improve biochemical monitoring in primary care, which reflects the increasing use of information technology in healthcare. Health information technology (HIT), including electronic prescribing and clinical decision support, has been shown to improve patient safety and reduce medication errors (Bates *et al.*, 1999; Garg *et al.*, 2005). In general, primary care is highly computerized and was an early adopter of HIT (Bryan & Boren, 2008; Ludwick & Doucette, 2009). Within the UK, general practice has the highest level of computer use and literacy in the NHS (BMA General Practitioners Committee, 2010). Therefore, there exists the necessary electronic infrastructure in which to introduce electronic alerts to remind the prescriber to undertake laboratory monitoring for patients treated with specific drugs. Indeed, when respondents to an electronic Delphi survey were asked to form a consensus on the most important features of GP computer systems for the improvement of patient safety, all agreed on the importance that 'it should be possible to set up the [computer] system so that patients can be automatically recalled for blood tests and other forms of monitoring' (Avery *et al.*, 2005).

Two systematic reviews have synthesized the literature on the use of HIT interventions to improve laboratory monitoring in primary care (Hayward *et al.*, 2009; Fischer *et al.*, 2010). Both reviews identified the same studies, which used various study designs and interventions. The systematic reviews both concluded that the literature demonstrated conflicting results. Hayward and colleagues (2009) stated that the studies that incorporated an intervention at a point in time outside of the GP visit (e.g. involvement of a clinical pharmacist team to arrange for patients to have a

laboratory test) led to increased monitoring, while the other health technology interventions (e.g. if monitoring was due, the electronic prescribing system generated a non-interruptive alert to recommend a laboratory test when a drug was prescribed) showed no improvement in laboratory monitoring. Fischer and colleagues (2010) summarized that although there was some improvement following the interventions to improve laboratory monitoring, when the analysis was restricted to the well-designed, higher quality studies, little improvement was seen with HIT interventions that only targetted doctors.

The use of electronic alerts and reminders targetted at GPs using information technology in daily practice may go some way towards improving laboratory monitoring, but there is the possibility that the problem of alert fatigue may reduce the effect of the reminders. Alert fatigue can occur when the alerts recommending monitoring do not relate to serious enough outcomes, are irrelevant, or because a given alert appears repeatedly and can therefore lead to the alert being overridden or simply ignored (Shah *et al.*, 2006; van der Sijs *et al.*, 2006; van der Sijs *et al.*, 2008; Isaac *et al.*, 2009).

There have been no efficient developments in electronic systems to alert GPs when a patient has not had a laboratory test that the doctor had recommended. Goldman (2010) presented results of interviews of doctors who described their annoyance that there were no efficient electronic systems in place to track whether patients had fulfilled the laboratory tests that had been ordered. This desire to have an electronic system to assist in the tracking of biochemical tests to completion has also been echoed by doctors in the secondary care setting (Poon *et al.*, 2004).

9.5.3 Patient factors

No work has been undertaken to investigate patients' perceptions on laboratory monitoring. There is a need for this type of research in order to identify other potential barriers to follow-up monitoring. Certainly the reason why such a large proportion of patients did not have follow-up monitoring following the initiation of their antihypertensive treatment may have been due to patients failing to attend for laboratory testing and not because GPs were failing to order the tests. Goldman and colleagues (2010) suggested that further work should investigate how doctors communicate the need for laboratory monitoring to patients and how patients perceive the role of monitoring in the detection and potential prevention of harm. As described earlier, interventions to improve laboratory monitoring using information technology have focused primarily on addressing issues in the work systems used by GPs. There is some evidence that interventions which involve the patients in the monitoring process (e.g. through the use of automated phone calls to patients due for drug monitoring) show significant improvement in the frequency of follow-up monitoring (Feldstein *et al.*, 2006). Future work should also be carried out to investigate the impact of interventions focussed specifically on improving the awareness of the importance of recommended monitoring in patients.

9.5.4 Uncertain responsibility for monitoring

Finally, an uncertain responsibility for monitoring has also been identified as a potential barrier to monitoring (Goldman *et al.*, 2010). Most patients with hypertension, especially those with 'simple' hypertension, are treated in the community and the general practitioner who initiated treatment would usually be

responsible for monitoring. However, in instances where treatment is initiated by specialist prescribers outside of primary care, it may be difficult to determine where the responsibility for monitoring lies.

9.6 Evaluating the cost-effectiveness of monitoring

Monitoring of patients for adverse reactions to treatment should be cost proportionate, in that the cost of testing should not be greater than the cost savings associated with reducing the health burden of adverse reactions to therapy. Doctors working in the American fee-for-service setting have stated that monitoring 'requires an inordinate amount of unreimbursed time (Goldman *et al.*, 2010).

Little to no work had been undertaken to evaluate the cost-effectiveness or cost-benefit of laboratory monitoring. Some research has been undertaken to determine the cost-effectiveness of therapeutic drug monitoring, which differs from monitoring for adverse reactions to treatment. The aim of therapeutic drug monitoring is to maximise therapeutic efficacy and prevent patient harm due to treatment with drugs with a narrow therapeutic range. This is achieved through regular blood measurements in order to maintain a relatively constant concentration of the drug in the bloodstream. Therapeutic drug monitoring has been demonstrated to be cost-effective for aminoglycosides and to a lesser extent for vancomycin, anti-epileptics, and immunosuppressant therapy (Touw *et al.*, 2005).

Only one study has focused specifically on the cost-effectiveness of laboratory monitoring of patients treated with antihypertensive therapy. The recent study by Smith and colleagues (Smith *et al.*, 2011) used a probabilistic decision model to

compare the cost of laboratory monitoring programme for patients treated with ACE inhibitors or AT-II receptor antagonists with the cost of the adverse outcomes of hyperkalaemia and acute renal failure. The programme involved a telephone call, followed by a letter, to a patient if they had not had a test of serum creatinine and potassium within five days of starting treatment. In the whole patient population, the per patient cost of adverse events was $119 with a monitoring programme, compared with $94 without the programme. Therefore, on average, the programme cost was $24 extra per person, per year. Cost savings were only observed when the analysis was restricted to the monitoring programme was targetted to patients with chronic kidney disease – patients who are higher risk of adverse events. Although the results are not generalizable outside of the system within which the analysis was undertaken—an American HMO with an established electronic health record system—the results do suggest that for laboratory monitoring of patients treated with antihypertensive therapy to be cost-effective, the monitoring needs to be carried out in a population with a high risk of adverse events.

9.7 Concluding observations

Monitoring of drug treatment can lead to better selection of drug therapy, better titration of treatment, improved adherence, and perhaps most importantly—identify potential adverse reactions to treatment before causing harm. I used a retrospective analysis of a large primary care database to examine the nature of biochemical monitoring in a cohort of patients newly treated with antihypertensive therapy. Although monitoring of renal function and electrolyte concentrations is recommended in published guidance, only half of patients newly treated with antihypertensive therapy had any follow-up monitoring within one year of treatment.

The benefits of monitoring as practiced in this cohort were slight to non-existent. Indeed, when monitoring was undertaken it was associated with increased risk of drug discontinuation, hospital admission and even death, although it is likely that this association was only a reflection of the clinician's willingness to reactively monitor patients who were already at greater risk of these events or who presented with symptoms requiring hospital admission or discontinuation of therapy.

The GPRD was able to show how monitoring was undertaken in primary care, but it could not demonstrate or develop the ideal monitoring strategy for adverse reactions to antihypertensive treatment. Indeed, there is a significant lack of primary evidence upon which to base rational monitoring strategies, which can only be addressed through the use of prospective observational or clinical trials of monitoring.

Several barriers to monitoring have been identified including the lack of systems in place to remind GPs to undertake monitoring. There is also a limited understanding of non-adherence to monitoring by patients, as the lack of monitoring may be as a result of the patient not attending for laboratory testing and not because the clinician did not order a test. More work needs to be carried out in order to determine patients' perceptions on the need for monitoring and research needs to be undertaken to evaluate the impact of interventions directed specifically to patients on the uptake of monitoring. Finally, more work is needed to understand the cost-effectiveness of monitoring by comparing the cost of monitoring with the cost of harm caused by the adverse reactions to antihypertensive treatment.

Future studies are necessary in order to determine whether more frequent or more assiduous biochemical monitoring in hypertensive patients in general practice would reduce harm from adverse drug reactions. Results from such studies would help to devise rational, evidence-based monitoring strategies.

References

Abayomi, K., Gelman, A. and Levy, M. (2008) Diagnostics for multivariate imputations. **Journal of the Royal Statistical Society: Series C Applied Statistics,** 57 (Part 3)**:** 273–291.

Abookire, S. A., Karson, A. S., Fiskio, J., et al. (2001) Use and monitoring of "statin" lipid-lowering drugs compared with guidelines. **Archives of Internal Medicine,** 161 (1)**:** 53–58.

Allison, P. D. (2009) "Missing data." In Millsap, R. E. and Maydeau-Olivares, A. (eds.) **The SAGE Handbook of Quantitative Methods in Psychology.** Thousand Oaks, CA: Sage Publications Inc. pp.72–89.

Andrade, S. E., Kahler, K. H., Frech, F., et al. (2006) Methods for evaluation of medication adherence and persistence using automated databases. **Pharmacoepidemiology and Drug Safety,** 15 (8)**:** 565–574.

Aronson, J. K. (2006) "Diuretics." In Aronson, J. K. (ed.) **Meyler's Side Effects of Drugs The International Encyclopedia of Adverse Drug Reactions and Interactions.** 15th ed. Amsterdam: Elsevier. pp.1152–1169.

Aronson, J. K. and Ferner, R. E. (2003) Joining the DoTS: new approach to classifying adverse drug reactions. **BMJ,** 327 (7425)**:** 1222–1225.

Ashraf, N., Locksley, R. and Arieff, A. I. (1981) Thiazide-induced hyponatremia associated with death or neurologic damage in outpatients. **American Journal of Medicine,** 70 (6)**:** 1163–1168.

Ashworth, M., Medina, J. and Morgan, M. (2008) Effect of social deprivation on blood pressure monitoring and control in England: a survey of data from the quality and outcomes framework. **BMJ,** 337**:** a2030.

Austin, P. C. (2007) Propensity-score matching in the cardiovascular surgery literature from 2004 to 2006: a systematic review and suggestions for improvement. **Journal of Thoracic and Cardiovascular Surgery,** 134 (5)**:** 1128–1135.

Austin, P. C. (2008a) Assessing balance in measured baseline covariates when using many-to-one matching on the propensity-score. **Pharmacoepidemiology and Drug Safety,** 17 (12)**:** 1218–1225.

Austin, P. C. (2008b) A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. **Statistics in Medicine,** 27 (12)**:** 2037–2049.

Austin, P. C. (2008c) Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score. **Pharmacoepidemiology and Drug Safety,** 17 (12)**:** 1202–1217.

Austin, P. C. (2008d) The performance of different propensity-score methods for estimating relative risks. **Journal of Clinical Epidemiology,** 61 (6)**:** 537–545.

Austin, P. C. (2008e) Primer on statistical interpretation or methods report card on propensity-score matching in the cardiology literature from 2004 to 2006: a systematic review. **Circulation: Cardiovascular Quality and Outcomes,** 1 (1)**:** 62–67.

Austin, P. C. (2009a) Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. **Statistics in Medicine,** 28 (25)**:** 3083–3107.

Austin, P. C. (2009b) The relative ability of different propensity score methods to balance measured covariates between treated and untreated subjects in observational studies. **Medical Decision Making,** 29 (6)**:** 661–677.

Austin, P. C. (2009c) Some methods of propensity-score matching had superior performance to others: results of an empirical investigation and Monte Carlo simulations. **Biometrical Journal,** 51 (1)**:** 171–184.

Austin, P. C., Grootendorst, P. and Anderson, G. M. (2007) A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. **Statistics in Medicine,** 26 (4)**:** 734–753.

Austin, P. C. and Mamdani, M. M. (2006) A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. **Statistics in Medicine,** 25 (12)**:** 2084–2106.

Austin, P. C., Mamdani, M. M., van Walraven, C., et al. (2006) Quantifying the impact of survivor treatment bias in observational studies. **Journal of Evaluation in Clinical Practice,** 12 (6)**:** 601–612.

Austin, P. C. and Platt, R. W. (2010) Survivor treatment bias, treatment selection bias, and propensity scores in observational research. **Journal of Clinical Epidemiology,** 63 (2)**:** 136–138.

Australian Heart Foundation (2008) **Guide to management of hypertension 2008 - assessing and managing raised blood pressure in adults** [online]. National Heart Foundation of Australia. http://www.heartfoundation.org.au/SiteCollectionDocuments/A%20Hypert%20Guidelines2008%20Guideline.pdf [Accessed 15 July 2010].

Avery, A. J., Savelyich, B. S., Sheikh, A., et al. (2005) Identifying and establishing consensus on the most important safety features of GP computer systems: e-Delphi study. **Informatics in Primary Care,** 13 (1)**:** 3–12.

Ayanian, J. Z., Landrum, M. B., Guadagnoli, E., et al. (2002) Specialty of ambulatory care physicians and mortality among elderly patients after myocardial infarction. **New England Journal of Medicine,** 347 (21)**:** 1678–1686.

Barzi, F. and Woodward, M. (2004) Imputations of missing values in practice: results from imputations of serum cholesterol in 28 cohort studies. **American Journal of Epidemiology,** 160 (1)**:** 34–45.

Baser, O. (2006) Too much ado about propensity score models? Comparing methods of propensity score matching. **Value Health,** 9 (6)**:** 377–385.

Bates, D. W., Teich, J. M., Lee, J., et al. (1999) The impact of computerized physician order entry on medication error prevention. **Journal of the American Medical Informatics Association,** 6 (4)**:** 313–321.

Bell, K. J., Hayen, A., Macaskill, P., et al. (2010) Monitoring initial response to angiotensin-converting enzyme inhibitor-based regimens: an individual patient data meta-analysis from randomized, placebo-controlled trials. **Hypertension,** 56 (3)**:** 533–539.

Bell, K. J., Hayen, A., Macaskill, P., et al. (2009) Value of routine monitoring of bone mineral density after starting bisphosphonate treatment: secondary analysis of trial data. **BMJ,** 338**:** b2266.

Ben Salem, C., Hmouda, H. and Bouraoui, K. (2009) Drug-induced hypokalaemia. **Current Drug Safety,** 4 (1)**:** 55–61.

Besançon, J. F., Lagarce, L., Diquet, B., et al. (2008) Study of the use of a spironolactone and angiotensin-converting enzyme inhibitor combination: a population-based analysis. **Pharmacoepidemiology and Drug Safety,** 17 (2)**:** 172–179.

Beyersmann, J., Gastmeier, P., Wolkewitz, M., et al. (2008a) An easy mathematical proof showed that time-dependent bias inevitably leads to biased effect estimation. **Journal of Clinical Epidemiology,** 61 (12)**:** 1216–1221.

Beyersmann, J., Wolkewitz, M. and Schumacher, M. (2008b) The impact of time-dependent bias in proportional hazards modelling. **Statistics in Medicine,** 27 (30)**:** 6439–6454.

Black, N. (1996) Why we need observational studies to evaluate the effectiveness of health care. **BMJ,** 312 (7040)**:** 1215–1218.

BMA General Practitioners Committee (2010) **Fit for the future - the evolution of general practice** [online]. http://www.bma.org.uk/images/fitforthefuture_tcm41-196103.pdf [Accessed 30 March 2011].

Bodner, T. E. (2008) What improves with increased missing data imputations? **Structural Equation Modeling: A Multidisciplinary Journal,** 15 (4)**:** 651–675.

Bourgault, C., Senecal, M., Brisson, M., et al. (2005) Persistence and discontinuation patterns of antihypertensive therapy among newly treated patients: a population-based study. **Journal of Human Hypertension,** 19 (8)**:** 607–613.

Braitman, L. E. and Rosenbaum, P. R. (2002) Rare outcomes, common treatments: analytic strategies using propensity scores. **Annals of Internal Medicine,** 137 (8)**:** 693–695.

Brand, J. P. L. (1999) **Development, implementation, and evaluation of multiple imputation strategies for the statistical analysis of incomplete data sets**. PhD thesis, Erasmus University.

British Columbia Ministry of Health Services (2008) **Hypertension: detection, diagnosis and management** [online]. http://www.bcguidelines.ca/pdf/hypertension_summary.pdf [Accessed 2 April 2010].

Brookhart, M. A., Schneeweiss, S., Rothman, K. J., et al. (2006) Variable selection for propensity score models. **American Journal of Epidemiology,** 163 (12)**:** 1149–1156.

Bryan, C. and Boren, S. A. (2008) The use and effectiveness of electronic clinical decision support tools in the ambulatory/primary care setting: a systematic review of the literature. **Informatics in Primary Care,** 16 (2)**:** 79–91.

Burger, D., Hugen, P., Reiss, P., et al. (2003) Therapeutic drug monitoring of nelfinavir and indinavir in treatment-naive HIV-1-infected individuals. **AIDS,** 17 (8)**:** 1157-1165.

Burke, T. A., Sturkenboom, M. C., Lu, S. E., et al. (2006) Discontinuation of antihypertensive drugs among newly diagnosed hypertensive patients in UK general practice. **Journal of Hypertension,** 24 (6)**:** 1193–1200.

Caetano, P. A., Lam, J. M. and Morgan, S. G. (2006) Toward a standard definition and measurement of persistence with drug therapy: Examples from research on statin and antihypertensive utilization. **Clinical Therapeutics,** 28 (9)**:** 1411–1424; discussion 1410.

Campbell, S. M., Roland, M. O., Middleton, E., et al. (2005) Improvements in quality of clinical care in English general practice 1998-2003: longitudinal observational study. **BMJ,** 331 (7525)**:** 1121.

Canadian Hypertension Education Program (2009) **2009 CHEP Recommendations for the Management of Hypertension** [online]. http://hypertension.ca/chep/wp-content/uploads/2009/01/2009-chep-recommendations_d3.pdf [Accessed 15 July 2010].

Carlin, J. B., Galati, J. C. and Royston, P. (2008) A new framework for managing and analyzing multiply imputed data in Stata. **Stata Journal,** 8 (1)**:** 49–67.

Cepeda, M. S., Boston, R., Farrar, J. T., et al. (2003) Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. **American Journal of Epidemiology,** 158 (3)**:** 280–287.

Chiu, T. F., Bullard, M. J., Chen, J. C., et al. (1997) Rapid life-threatening hyperkalemia after addition of amiloride HCl/hydrochlorothiazide to angiotensin-converting enzyme inhibitor therapy. **Annals of Emergency Medicine,** 30 (5)**:** 612–615.

Chobanian, A. V., Bakris, G. L., Black, H. R., et al. (2003) Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. **Hypertension,** 42 (6)**:** 1206–1252.

Chow, K. M., Szeto, C. C., Wong, T. Y., et al. (2003) Risk factors for thiazide-induced hyponatraemia. **QJM,** 96 (12)**:** 911–917.

Clark, B. A., Shannon, R. P., Rosa, R. M., et al. (1994) Increased susceptibility to thiazide-induced hyponatremia in the elderly. **Journal of the American Society of Nephrology,** 5 (4)**:** 1106–1111.

Clark, D. A., Stinson, E. B., Griepp, R. B., et al. (1971) Cardiac transplantation in man. VI. Prognosis of patients selected for cardiac transplantation. **Annals of Internal Medicine,** 75 (1)**:** 15–21.

Clayton, J. A., Le Jeune, I. R. and Hall, I. P. (2006a) Severe hyponatraemia in medical in-patients: aetiology, assessment and outcome. **QJM,** 99 (8)**:** 505–511.

Clayton, J. A., Rodgers, S., Blakey, J., et al. (2006b) Thiazide diuretic prescription and electrolyte abnormalities in primary care. **British Journal of Clinical Pharmacology,** 61 (1)**:** 87–95.

Cochran, W. G. and Rubin, D. B. (1973) Controlling bias in observational research: a review. **Sankhya Series A,** 35**:** 417–446.

Cole, J. A., Loughlin, J. E., Ajene, A. N., et al. (2002) The effect of zanamivir treatment on influenza complications: a retrospective cohort study. **Clinical Therapeutics,** 24 (11)**:** 1824–1839.

Coleman, J. J., Ferner, R. E. and Evans, S. J. (2006) Monitoring for adverse drug reactions. **British Journal of Clinical Pharmacology,** 61 (4)**:** 371–378.

Coleman, J. J., McDowell, S. E. and Ferner, R. E. (2008) Where does the evidence for guidelines come from? An example of monitoring for the adverse renal effects of ACE inhibitors. **Drug Safety,** 31 (10)**:** 905–906.

Corry, M., Bonner, G., McEntee, S., et al. (2000) Hospitals do not inform GPs about medication that should be monitored. **Family Practice,** 17 (3)**:** 268–271.

Cox, D. R. (1972) Regression models and life-tables. **Journal of the Royal Statistical Society: Series B Statistical Methodology,** 34 (2)**:** 187–220.

Cox, D. R. and Oakes, D. (2001) **Analysis of survival data**. London: Chapman and Hall.

D'Agostino, R. B., Jr. (1998) Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. **Statistics in Medicine,** 17 (19)**:** 2265–2281.

Davis, S., King, B. and Raine, J. (2007) "Spontaneous Reporting - UK." In Mann, R. D. and Andrews, E. B. (eds.) **Pharmacovigilance.** 2nd ed. London: John Wiley & Sons Ltd. pp.199–215.

Deeks, J. J., Dinnes, J., D'Amico, R., et al. (2003) Evaluating non-randomised intervention studies. **Health Technology Assessment,** 7 (27)**:** iii–x, 1–173.

Dehejia, R. H. and Wahba, S. (2002) Propensity-score matching methods for nonexperimental causal studies. **The Review of Economics and Statistics,** 84 (1)**:** 151–161.

Delaney, J. A., Daskalopoulou, S. S., Brophy, J. M., et al. (2007) Lifestyle variables and the risk of myocardial infarction in the general practice research database. **BMC Cardiovascular Disorders,** 7**:** 38.

Delaney, J. A., Daskalopoulou, S. S. and Suissa, S. (2009) Traditional versus marginal structural models to estimate the effectiveness of beta-blocker use on mortality after myocardial infarction. **Pharmacoepidemiology and Drug Safety,** 18 (1)**:** 1–6.

Delaney, J. A., Moodie, E. E. and Suissa, S. (2008) Validating the effects of drug treatment on blood pressure in the General Practice Research Database. **Pharmacoepidemiology and Drug Safety,** 17 (6)**:** 535–545.

Department of Veterans Affairs/ Department of Defense (2007) **VA/DoD Clinical Practice Guideline for Management of Chronic Kidney Disease in Primary Care** [online]. http://www.healthquality.va.gov/ckd/ckd_v478.pdf [Accessed 2 April 2010].

Donders, A. R., van der Heijden, G. J., Stijnen, T., et al. (2006) Review: a gentle introduction to imputation of missing values. **Journal of Clinical Epidemiology,** 59 (10)**:** 1087–1091.

Dorup, I., Skajaa, K., Clausen, T., et al. (1988) Reduced concentrations of potassium, magnesium, and sodium-potassium pumps in human skeletal muscle during treatment with diuretics. **British Medical Journal (Clinical Research Ed),** 296 (6620)**:** 455–458.

Ebbesen, J., Buajordet, I., Erikssen, J., et al. (2001) Drug-related deaths in a department of internal medicine. **Archives of Internal Medicine,** 161 (19)**:** 2317–2323.

Eccles, M., Freemantle, N. and Mason, J. (1998) North of England evidence based development project: guideline for angiotensin converting enzyme inhibitors in primary care management of adults with symptomatic heart failure. **BMJ,** 316 (7141)**:** 1369–1375.

Elliott, W. J., Plauschinat, C. A., Skrepnek, G. H., et al. (2007) Persistence, adherence, and risk of discontinuation associated with commonly prescribed antihypertensive drug monotherapies. **Journal of the American Board of Family Medicine,** 20 (1)**:** 72–80.

Feldstein, A. C., Smith, D. H., Perrin, N., et al. (2006) Improved therapeutic monitoring with several interventions: a randomized trial. **Archives of Internal Medicine,** 166 (17)**:** 1848–1854.

Ferner, R. E., Coleman, J., Pirmohamed, M., et al. (2005) The quality of information on monitoring for haematological adverse drug reactions. **British Journal of Clinical Pharmacology,** 60 (4)**:** 448–451.

Fichman, M. P., Vorherr, H., Kleeman, C. R., et al. (1971) Diuretic-induced hyponatremia. **Annals of Internal Medicine,** 75 (6)**:** 853–863.

Fischer, S. H., Tjia, J. and Field, T. S. (2010) Impact of health information technology interventions to improve medication laboratory monitoring for ambulatory patients: a systematic review. **Journal of the American Medical Informatics Association,** 17 (6)**:** 631–636.

Franse, L. V., Pahor, M., Di Bari, M., et al. (2000) Hypokalemia associated with diuretic use and cardiovascular events in the Systolic Hypertension in the Elderly Program. **Hypertension,** 35 (5)**:** 1025–1030.

Freis, E. D. (1995) The efficacy and safety of diuretics in treating hypertension. **Annals of Internal Medicine,** 122 (3)**:** 223–226.

French Haute Autorité de Santé (2005) **Management of adults with essential hypertension** [online]. http://www.has-sante.fr/portail/upload/docs/application/pdf/Hypertension_guidelines.pdf [Accessed 28 Feburary 2010].

Gail, M. H. (1972) Does cardiac transplantation prolong life? A reassessment. **Annals of Internal Medicine,** 76 (5)**:** 815–817.

Gandhi, T. K., Weingart, S. N., Borus, J., et al. (2003) Adverse drug events in ambulatory care. **New England Journal of Medicine,** 348 (16)**:** 1556–1564.

Garcia Rodriguez, L. A. and Perez Gutthann, S. (1998) Use of the UK General Practice Research Database for pharmacoepidemiology. **British Journal of Clinical Pharmacology,** 45 (5)**:** 419–425.

Garg, A. X., Adhikari, N. K., McDonald, H., et al. (2005) Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. **JAMA,** 293 (10)**:** 1223–1238.

Gayat, E., Pirracchio, R., Resche-Rigon, M., et al. (2010) Propensity scores in intensive care and anaesthesiology literature: a systematic review. **Intensive Care Medicine**.

Gelfand, J. M., Margolis, D. J. and Dattani, H. (2005) "The UK General Practice Research Database." In Strom, B. L. (ed.) **Pharmacoepidemiology.** 4th ed. Chichester: John Wiley & Sons. pp.337–346.

General Practice Research Database (2007) **GPRD - an overview**. London: Crown Copyright.

Gennari, F. J. (1998) Hypokalemia. **New England Journal of Medicine,** 339 (7)**:** 451–458.

Gérardin-Marais, M., Victorri-Vigneau, C., Allain-Veyrac, G., et al. (2008) Diuretic drug therapy monitoring in the elderly: a cohort study. **European Journal of Clinical Pharmacology,** 64 (4)**:** 433–437.

Glasziou, P. (2007) How much monitoring? **British Journal of General Practice,** 57 (538)**:** 350–351.

Glasziou, P. and Aronson, J. K. (2008) "An introduction to monitoring therapeutic interventions in clinical practice." In Glasziou, P., Irwig, L. and Aronson, J. K. (eds.) **Evidence-based medical monitoring: from principles to practice.** Oxford: Blackwell Publishing Ltd. pp.3–14.

Glasziou, P., Irwig, L. and Mant, D. (2005) Monitoring in chronic disease: a rational approach. **BMJ,** 330 (7492)**:** 644–648.

Glasziou, P. P., Irwig, L., Heritier, S., et al. (2008) Monitoring cholesterol levels: measurement error or true change? **Annals of Internal Medicine,** 148 (9)**:** 656–661.

Glesby, M. J. and Hoover, D. R. (1996) Survivor treatment selection bias in observational studies: examples from the AIDS literature. **Annals of Internal Medicine,** 124 (11)**:** 999–1005.

Glynn, R. J., Schneeweiss, S. and Stürmer, T. (2006) Indications for propensity scores and review of their use in pharmacoepidemiology. **Basic Clinical Pharmacology & Toxicology,** 98 (3)**:** 253–259.

Goldman, R. E., Soran, C. S., Hayward, G. L., et al. (2010) Doctors' perceptions of laboratory monitoring in office practice. **Journal of Evaluation in Clinical Practice,** 16 (6)**:** 1136–1141.

Gorelick, M. H. (2006) Bias arising from missing data in predictive models. **Journal of Clinical Epidemiology,** 59 (10)**:** 1115–1123.

Graham, D. J., Drinkard, C. R., Shatin, D., et al. (2001) Liver enzyme monitoring in patients treated with troglitazone. **JAMA,** 286 (7)**:** 831–833.

Greenland, S. and Finkle, W. D. (1995) A critical look at methods for handling missing covariates in epidemiologic regression analyses. **American Journal of Epidemiology,** 142 (12)**:** 1255–1264.

Greenland, S. and Morgenstern, H. (2001) Confounding in health research. **Annual Review of Public Health,** 22**:** 189–212.

Greevy, R. A., Huizinga, M. M., Roumie, C. L., et al. (2010) The impact of stockpiling and minimum required adherence on estimating persistence on oral anti-diabetic therapy from electronic health records. **26th Annual Conference on Pharmacoepidemiology & Therapeutic Risk Management. Brighton, UK, 19–22 August 2010**.

Gross, P. and Palm, C. (2005) Thiazides: do they kill? **Nephrology, Dialysis, Transplantation,** 20 (11)**:** 2299–2301.

Guidelines Committee (2003) 2003 European Society of Hypertension-European Society of Cardiology guidelines for the management of arterial hypertension. **Journal of Hypertension,** 21 (6)**:** 1011–1053.

Gurwitz, J. H., Field, T. S., Harrold, L. R., et al. (2003) Incidence and preventability of adverse drug events among older persons in the ambulatory setting. **JAMA,** 289 (9)**:** 1107–1116.

Hardy, J. R., Holford, T. R., Hall, G. C., et al. (2004) Strategies for identifying pregnancies in the automated medical records of the General Practice Research Database. **Pharmacoepidemiology and Drug Safety,** 13 (11)**:** 749–759.

Harel, O. and Zhou, X. H. (2007) Multiple imputation: review of theory, implementation and software. **Statistics in Medicine,** 26 (16)**:** 3057–3077.

Harrell, F. E., Jr., Lee, K. L. and Mark, D. B. (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. **Statistics in Medicine,** 15 (4)**:** 361–387.

Hasford, J., Mimran, A. and Simons, W. R. (2002) A population-based European cohort study of persistence in newly diagnosed hypertensive patients. **Journal of Human Hypertension,** 16 (8)**:** 569–575.

Hayward, G. L., Parnes, A. J. and Simon, S. R. (2009) Using health information technology to improve drug monitoring: a systematic review. **Pharmacoepidemiology and Drug Safety,** 18 (12)**:** 1232–1237.

Hebert, P. L. (2010) **Combining propensity scores with instrumental variable analysis** [online]. http://www.academyhealth.org/files/2010/tuesday/HebertP.pdf [Accessed 30 August 2010].

Heckman, J., Ichimura, H., Smith, J., et al. (1998) Characterizing selection bias using experimental data. **Econometrica,** 66 (5)**:** 1017–1098.

Hennekens, C. H. and Buring, J. E. (1987) **Epidemiology in medicine**. Boston: Little, Brown.

Herrett, E., Thomas, S. L., Schoonen, W. M., et al. (2010) Validation and validity of diagnoses in the General Practice Research Database: a systematic review. **British Journal of Clinical Pharmacology,** 69 (1)**:** 4–14.

Hill, J. (2004) **Reducing bias in treatment effect estimation in observational studies suffering from missing data. ISERP Working Paper 04-01**. New York: Columbia University Institute for Social and Economic Research and Policy.

Hill, J. L., Reiter, J. P. and Zanutto, E. (2004) "A comparison of experimental and observational data analyses." In Gelman, A. and Meng, X. (eds.) **Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives.** New York: John Wiley & Sons. pp.49–60.

Ho, D. E., Imai, K., King, G., et al. (2007) Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. **Political Analysis,** 15**:** 199–236.

Hoch, I., Heymann, A. D., Kurman, I., et al. (2003) Countrywide computer alerts to community physicians improve potassium testing in patients receiving diuretics. **Journal of the American Medical Informatics Association,** 10 (6)**:** 541–546.

Hoffmann, F. and Andersohn, F. (2011) Immortal time bias and survival in patients who self-monitor blood glucose in the Retrolective Study: self-monitoring of Blood Glucose and Outcome in Patients with Type 2 Diabetes (ROSSO). **Diabetologia,** 54 (2)**:** 308–311.

Horton, N. J. and Kleinman, K. P. (2007) Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. **The American Statistician,** 61 (1)**:** 79–90.

Hosmer, D. W., Hosmer, T., Le Cessie, S., et al. (1997) A comparison of goodness-of-fit tests for the logistic regression model. **Statistics in Medicine,** 16 (9)**:** 965–980.

Howard, R. L., Avery, A. J., Howard, P. D., et al. (2003) Investigation into the reasons for preventable drug related admissions to a medical admissions unit: observational study. **Quality & Safety in Health Care,** 12 (4)**:** 280–285.

Hurley, J. S., Roberts, M., Solberg, L. I., et al. (2005) Laboratory safety monitoring of chronic medications in ambulatory care settings. **Journal of General Internal Medicine,** 20 (4)**:** 331–333.

Imai, K., King, G. and Stuart, E. A. (2008) Misunderstandings between experimentalists and observationalists about causal inference. **Journal of the Royal Statistical Society: Series A,** 171 (Part 2)**:** 481–502.

Indermitte, J., Burkolter, S., Drewe, J., et al. (2007) Risk factors associated with a high velocity of the development of hyperkalaemia in hospitalised patients. **Drug Safety,** 30 (1)**:** 71–80.

Institute for Clinical Systems Improvement (2010) **Health Care Guideline: Hypertension Diagnosis and Treatment** [online]. http://www.icsi.org/hypertension_4/hypertension_diagnosis_and_treatment_4.html [Accessed 2 April 2010].

Isaac, T., Weissman, J. S., Davis, R. B., et al. (2009) Overrides of medication alerts in ambulatory care. **Archives of Internal Medicine,** 169 (3)**:** 305–311.

Jannuzzi, G., Cian, P., Fattore, C., et al. (2000) A multicenter randomized controlled trial on the clinical impact of therapeutic drug monitoring in patients with newly diagnosed epilepsy. The Italian TDM Study Group in Epilepsy. **Epilepsia,** 41 (2)**:** 222-230.

Jiang, J. Y., Wong, M. C., Ali, M. K., et al. (2009) Association of antihypertensive monotherapy with serum sodium and potassium levels in Chinese patients. **American Journal of Hypertension,** 22 (3)**:** 243–249.

Jick, H., Derby, L. E., Gurewich, V., et al. (1996) The risk of myocardial infarction associated with antihypertensive drug treatment in persons with uncomplicated essential hypertension. **Pharmacotherapy,** 16 (3)**:** 321–326.

Jick, H., Jick, S. S. and Derby, L. E. (1991) Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. **BMJ,** 302 (6779)**:** 766–768.

Jick, H., Terris, B. Z., Derby, L. E., et al. (1992) Further validation of information recorded on a general practitioner based computerized data resource in the United Kingdom. **Pharmacoepidemiology and Drug Safety,** 1**:** 347–349.

Joint Formulary Committee (2010) **British National Formulary No.59**. London: BMJ Group and RPS Publishing.

Jones, M. P. (1996) Indicator and stratification methods for missing explanatory variables in multiple linear regression. **Journal of the American Statistical Association,** 91 (433)**:** 222–230.

Juntti-Patinen, L. and Neuvonen, P. J. (2002) Drug-related deaths in a university central hospital. **European Journal of Clinical Pharmacology,** 58 (7)**:** 479–482.

Juurlink, D. N., Mamdani, M. M., Lee, D. S., et al. (2004) Rates of hyperkalemia after publication of the Randomized Aldactone Evaluation Study. **New England Journal of Medicine,** 351 (6)**:** 543–551.

Kalra, P. A., Kumwenda, M., MacDowall, P., et al. (1999) Questionnaire study and audit of use of angiotensin converting enzyme inhibitor and monitoring in general practice: the need for guidelines to prevent renal failure. **BMJ,** 318 (7178)**:** 234–237.

Khan, N. F., Harrison, S. E. and Rose, P. W. (2010) Validity of diagnostic coding within the General Practice Research Database: a systematic review. **British Journal of General Practice,** 60 (572)**:** e128–136.

Knight, E. L. and Avorn, J. (2001) Quality indicators for appropriate medication use in vulnerable elders. **Annals of Internal Medicine,** 135 (8 Pt 2)**:** 703–710.

Kumar, A., Zarychanski, R., Light, B., et al. (2010) Early combination antibiotic therapy yields improved survival compared with monotherapy in septic shock: a propensity-matched analysis. **Critical Care Medicine,** 38 (9)**:** 1773–1785.

Kurth, T., Walker, A. M., Glynn, R. J., et al. (2006) Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. **American Journal of Epidemiology,** 163 (3)**:** 262–270.

Lafata, J. E., Gunter, M. J., Hsu, J., et al. (2007) Academic detailing to improve laboratory testing among outpatient medication users. **Medical Care,** 45 (10)**:** 966–972.

Lalani, T., Cabell, C. H., Benjamin, D. K., et al. (2010) Analysis of the impact of early surgery on in-hospital mortality of native valve endocarditis: use of propensity score and instrumental variable methods to adjust for treatment-selection bias. **Circulation,** 121 (8)**:** 1005–1013.

Lawrenson, R., Todd, J. C., Leydon, G. M., et al. (2000) Validation of the diagnosis of venous thromboembolism in general practice database studies. **British Journal of Clinical Pharmacology,** 49 (6)**:** 591–596.

Lawrenson, R., Williams, T. and Farmer, R. (1999) Clinical information for research; the use of general practice databases. **Journal of Public Health Medicine,** 21 (3)**:** 299–304.

Lee, K. J. and Carlin, J. B. (2010) Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. **American Journal of Epidemiology,** 171 (5)**:** 624–632.

Leuven, L. and Sianesi, B. (2003) **PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing**. http://ideas.repec.org/c/boc/bocode/s432001.html

Lévesque, L. E., Hanley, J. A., Kezouh, A., et al. (2010) Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. **BMJ,** 340**:** b5087.

Lewis, J. D., Brensinger, C., Bilker, W. B., et al. (2002) Validity and completeness of the General Practice Research Database for studies of inflammatory bowel disease. **Pharmacoepidemiology and Drug Safety,** 11 (3)**:** 211–218.

Liamis, G., Milionis, H. and Elisaf, M. (2008a) Blood pressure drug therapy and electrolyte disturbances. **International Journal of Clinical Practice,** 62 (10)**:** 1572–1580.

Liamis, G., Milionis, H. and Elisaf, M. (2008b) A review of drug-induced hyponatremia. **American Journal of Kidney Diseases,** 52 (1)**:** 144–153.

Lindenauer, P. K., Rothberg, M. B., Nathanson, B. H., et al. (2010) Activated protein C and hospital mortality in septic shock: a propensity-matched analysis. **Critical Care Medicine,** 38 (4)**:** 1101–1107.

Lloyd, S. J. and Mauro, V. F. (2000) Spironolactone in the treatment of congestive heart failure. **Annals of Pharmacotherapy,** 34 (11)**:** 1336–1340.

Lo, H. G., Matheny, M. E., Seger, D. L., et al. (2009) Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care. **Journal of the American Medical Informatics Association,** 16 (1)**:** 66–71.

Lohr, J. W. (1994) Osmotic demyelination syndrome following correction of hyponatremia: association with hypokalemia. **American Journal of Medicine,** 96 (5)**:** 408–413.

Ludwick, D. A. and Doucette, J. (2009) Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries. **International Journal of Medical Informatics,** 78 (1)**:** 22–31.

Lunt, M. (2008) **nscores: Transform to and from normal scores**. http://personalpages.manchester.ac.uk/staff/mark.lunt/nscores.html

Lunt, M., Solomon, D., Rothman, K., et al. (2009) Different methods of balancing covariates leading to different effect estimates in the presence of effect modification. **American Journal of Epidemiology,** 169 (7)**:** 909–917.

Mackay, A., Brown, J. J., Cumming, A. M., et al. (1979) Smoking and renal artery stenosis. **British Medical Journal,** 2 (6193)**:** 770.

Mancia, G., De Backer, G., Dominiczak, A., et al. (2007) 2007 Guidelines for the management of arterial hypertension: The Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). **European Heart Journal,** 28 (12)**:** 1462–1536.

Mann, S. J. (2008) The silent epidemic of thiazide-induced hyponatremia. **Journal of Clinical Hypertension (Greenwich, Conn),** 10 (6)**:** 477–484.

Mant, D. (2008) "A framework for developing and evaluating a monitoring strategy." In Glasziou, P., Irwig, L. and Aronson, J. K. (eds.) **Evidence-based medical monitoring: from principles to practice.** Oxford: Blackwell Publishing Ltd. pp.15–30.

Marston, L., Carpenter, J. R., Walters, K. R., et al. (2010) Issues in multiple imputation of missing data for large general practice clinical databases. **Pharmacoepidemiology and Drug Safety,** 19 (6)**:** 618–626.

Martens, E. P. (2007) **Methods to adjust for confounding. Propensity scores and instrumental variables.** PhD thesis, Utrecht University.

Martin, S., Schneider, B., Heinemann, L., et al. (2006) Self-monitoring of blood glucose in type 2 diabetes and long-term outcome: an epidemiological cohort study. **Diabetologia,** 49 (2)**:** 271–278.

Matheny, M. E., Sequist, T. D., Seger, A. C., et al. (2008) A randomized trial of electronic clinical reminders to improve medication laboratory monitoring. **Journal of the American Medical Informatics Association,** 15 (4)**:** 424–429.

McAlister, F. A., Tu, K., Majumdar, S. R., et al. (2007) Laboratory testing in newly treated elderly hypertensive patients without co-morbidities: a population-based cohort study. **Open Medicine,** 1 (2)**:** E60–67.

McAreavey, D., Brown, J. J., Cumming, A. M., et al. (1983) Inverse relation of exchangeable sodium and blood pressure in hypertensive patients with renal artery stenosis. **Journal of Hypertension,** 1 (3)**:** 297–302.

McMurray, J., Cohen-Solal, A., Dietz, R., et al. (2005) Practical recommendations for the use of ACE inhibitors, beta-blockers, aldosterone antagonists and angiotensin receptor blockers in heart failure: putting guidelines into practice. **European Journal of Heart Failure,** 7 (5)**:** 710–721.

Meier, C. R., Schlienger, R. G., Kraenzlin, M. E., et al. (2000) HMG-CoA reductase inhibitors and the risk of fractures. **JAMA,** 283 (24)**:** 3205–3210.

Messmer, B. J., Nora, J. J., Leachman, R. D., et al. (1969) Survival-times after cardiac allografts. **Lancet,** 1 (7602)**:** 954–956.

Michigan Quality Improvement Consortium (2009) **Medical Management of Adults with Hypertension** [online]. http://www.uphpplus.com/docs%5CHypertensionMQIC.pdf [Accessed 2 April 2010].

Miettinen, O. S. (1976) Stratification by a multivariate confounder score. **American Journal of Epidemiology,** 104 (6)**:** 609–620.

Millar, J. S. (2001) Consultations owing to adverse drug reactions in a single practice. **British Journal of General Practice,** 51 (463)**:** 130–131.

Miltiadous, G., Mikhailidis, D. P. and Elisaf, M. (2003) Acid-base and electrolyte abnormalities observed in patients receiving cardiovascular drugs. **Journal of Cardiovascular Pharmacology,** 8 (4)**:** 267–276.

Mjorndal, T., Boman, M. D., Hagg, S., et al. (2002) Adverse drug reactions as a cause for admissions to a department of internal medicine. **Pharmacoepidemiology and Drug Safety,** 11 (1)**:** 65–72.

Moons, K. G., Donders, R. A., Stijnen, T., et al. (2006) Using the outcome for imputation of missing predictor values was preferred. **Journal of Clinical Epidemiology,** 59 (10)**:** 1092–1101.

Moser, M. (1998) Why are physicians not prescribing diuretics more frequently in the management of hypertension? **JAMA,** 279 (22)**:** 1813–1816.

Moss, R. R., Humphries, K. H., Gao, M., et al. (2003) Outcome of mitral valve repair or replacement: a comparison by propensity score analysis. **Circulation,** 108 Suppl 1**:** II90–97.

National Collaborating Centre for Chronic Conditions (2006) **Hypertension: management in adults in primary care: pharmacological update**. London: Royal College of Physicians.

National Institute for Health and Clinical Excellence (2008) **NICE clinical guideline 73 - Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care** [online]. http://www.nice.org.uk/nicemedia/live/12069/42117/42117.pdf [Accessed 2 April 2010].

National Kidney Foundation (2002) **KDOQI Clinical Practice Guidelines on Hypertension and Antihypertensive Agents in Chronic Kidney Disease** [online]. http://www.kidney.org/professionals/KDOQI/guidelines_bp/index.htm [Accessed 2 April 2010].

NCQA (2010) **Annual monitoring for patients on persistent medications: measured for HEDIS 2010** [online]. www.fchp.org/providers/resources/~media/Files/FCHP/Imported/2010_HEDIS_Measures.pdf.ashx [Accessed 20 August 2010].

NHS Employers (2011) **Quality and outcomes framework** [online]. http://www.nhsemployers.org/PayAndContracts/GeneralMedicalServicesContract/QOF/Pages/QualityOutcomesFramework.aspx [Accessed 21 January 2011].

Nicholson, J. P., Teichman, S. L., Alderman, M. H., et al. (1983) Cigarette smoking and renovascular hypertension. **Lancet,** 2 (8353)**:** 765–766.

Normand, S. T., Landrum, M. B., Guadagnoli, E., et al. (2001) Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. **Journal of Clinical Epidemiology,** 54 (4)**:** 387–398.

North East Essex Medicines Management Committee (2007) **NEEMMC Guidelines on ACE Inhibitor (ACE-I) in Heart Failure** [online]. http://www.northeastessexpct.nhs.uk/guidelines/er-guidelines-for-heart-failure-june07.pdf [Accessed 8 March 2010].

Northern Ireland Department of Health Social Services and Public Safety (2007) **Pharmaceutical Clinical Effectiveness Programme - Supporting material for prescribing guideline for ACE inhibitors (and Angiotensin-II Receptor Antagonists)** [online]. http://www.dhsspsni.gov.uk/acei_notes_final.pdf [Accessed 28 Feburary 2010].

Nuttall, G. A. and Houle, T. T. (2008) Liars, damn liars, and propensity scores. **Anesthesiology,** 108 (1)**:** 3–4.

Ogihara, T., Kikuchi, K., Matsuoka, H., et al. (2009) The Japanese Society of Hypertension Guidelines for the Management of Hypertension (JSH 2009). **Hypertension Research,** 32 (1)**:** 3–107.

Olsen, H., Klemetsrud, T., Stokke, H. P., et al. (1999) Adverse drug reactions in current antihypertensive therapy: a general practice survey of 2586 patients in Norway. **Blood Pressure,** 8 (2)**:** 94–101.

Opatrny, L., Delaney, J. A. and Suissa, S. (2008) Gastro-intestinal haemorrhage risks of selective serotonin receptor antagonist therapy: a new look. **British Journal of Clinical Pharmacology,** 66 (1)**:** 76–81.

Osborn, D. P., Levy, G., Nazareth, I., et al. (2007) Relative risk of cardiovascular and cancer mortality in people with severe mental illness from the United Kingdom's General Practice Research Database. **Archives of General Psychiatry,** 64 (2)**:** 242–249.

Palen, T. E., Raebel, M., Lyons, E., et al. (2006) Evaluation of laboratory monitoring alerts within a computerized physician order entry system for medication orders. **American Journal of Managed Care,** 12 (7)**:** 389–395.

Palmer, B. F. (2003) Angiotensin-converting enzyme inhibitors and angiotensin receptor blockers: what to do if the serum creatinine and/or serum potassium concentration rises. **Nephrology, Dialysis, Transplantation,** 18 (10)**:** 1973–1975.

Palmer, B. F. (2004) Managing hyperkalemia caused by inhibitors of the renin-angiotensin-aldosterone system. **New England Journal of Medicine,** 351 (6)**:** 585–592.

Peduzzi, P., Concato, J., Feinstein, A. R., et al. (1995) Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. **Journal of Clinical Epidemiology,** 48 (12)**:** 1503–1510.

Peduzzi, P., Concato, J., Kemper, E., et al. (1996) A simulation study of the number of events per variable in logistic regression analysis. **Journal of Clinical Epidemiology,** 49 (12)**:** 1373–1379.

Perazella, M. A. (2000) Drug-induced hyperkalemia: old culprits and new offenders. **American Journal of Medicine,** 109 (4)**:** 307–314.

Perkins, S. M., Tu, W., Underhill, M. G., et al. (2000) The use of propensity scores in pharmacoepidemiologic research. **Pharmacoepidemiology and Drug Safety,** 9 (2)**:** 93–101.

Perreault, S., Lamarre, D., Blais, L., et al. (2005) Persistence with treatment in newly treated middle-aged patients with essential hypertension. **Annals of Pharmacotherapy,** 39 (9)**:** 1401–1408.

Pirmohamed, M. and Ferner, R. E. (2003) Monitoring drug treatment. **BMJ,** 327 (7425)**:** 1179–1181.

Pirmohamed, M., James, S., Meakin, S., et al. (2004) Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. **BMJ,** 329 (7456)**:** 15–19.

Poon, E. G., Gandhi, T. K., Sequist, T. D., et al. (2004) "I wish I had seen this test result earlier!": Dissatisfaction with test result management systems in primary care. **Archives of Internal Medicine,** 164 (20)**:** 2223–2228.

Posner, M. A. and Ash, A. S. (2006) **Comparing weighting methods in propensity score analysis** [online]. http://www.stat.columbia.edu/~gelman/stuff_for_blog/posner.pdf [Accessed 20 May 2010].

Pouyanne, P., Haramburu, F., Imbs, J. L., et al. (2000) Admissions to hospital caused by adverse drug reactions: cross sectional incidence study. French Pharmacovigilance Centres. **BMJ,** 320 (7241)**:** 1036.

Psaty, B. M., Koepsell, T. D., Yanez, N. D., et al. (1995) Temporal patterns of antihypertensive medication use among older adults, 1989 through 1992. An effect of the major clinical trials on clinical practice? **JAMA,** 273 (18)**:** 1436–1438.

Rademaker, M. (2001) Do women have more adverse drug reactions? **American Journal of Clinical Dermatology,** 2 (6)**:** 349–351.

Raebel, M. A., Lyons, E. E., Andrade, S. E., et al. (2005) Laboratory monitoring of drugs at initiation of therapy in ambulatory care. **Journal of General Internal Medicine,** 20 (12)**:** 1120–1126.

Raebel, M. A., McClure, D. L., Chan, K. A., et al. (2007a) Laboratory evaluation of potassium and creatinine among ambulatory patients prescribed spironolactone: are we monitoring for hyperkalemia? **Annals of Pharmacotherapy,** 41 (2)**:** 193–200.

Raebel, M. A., McClure, D. L., Simon, S. R., et al. (2007b) Laboratory monitoring of potassium and creatinine in ambulatory patients receiving angiotensin converting enzyme inhibitors and angiotensin receptor blockers. **Pharmacoepidemiology and Drug Safety,** 16 (1)**:** 55–64.

Raebel, M. A., McClure, D. L., Simon, S. R., et al. (2006) Frequency of serum creatinine monitoring during allopurinol therapy in ambulatory patients. **Annals of Pharmacotherapy,** 40 (3)**:** 386–391.

Raebel, M. A., Ross, C., Xu, S., et al. (2010) Diabetes and drug-associated hyperkalemia: effect of potassium monitoring. **Journal of General Internal Medicine,** 25 (4)**:** 326–333.

Ramadan, F. H., Masoodi, N. and El-Solh, A. A. (2005) Clinical factors associated with hyperkalemia in patients with congestive heart failure. **Journal of Clinical Pharmacy and Therapeutics,** 30 (3)**:** 233–239.

Ramsay, L., Williams, B., Johnston, G., et al. (1999) Guidelines for management of hypertension: report of the third working party of the British Hypertension Society. **Journal of Human Hypertension,** 13 (9)**:** 569–592.

Redelmeier, D. A. and Singh, S. M. (2001) Survival in Academy Award-winning actors and actresses. **Annals of Internal Medicine,** 134 (10)**:** 955–962.

Rhodes, K. E. (1992) Prescription of diuretic drugs and monitoring of long-term use in one general practice. **British Journal of General Practice,** 42 (355)**:** 68–70.

Rimmer, J. M., Horn, J. F. and Gennari, F. J. (1987) Hyperkalemia as a complication of drug therapy. **Archives of Internal Medicine,** 147 (5)**:** 867–869.

Rochon, P. A., Tu, J. V., Anderson, G. M., et al. (2000) Rate of heart failure and 1-year survival for older people receiving low-dose beta-blocker therapy after myocardial infarction. **Lancet,** 356 (9230)**:** 639–644.

Rosenbaum, P. R. and Rubin, D. B. (1983) The central role of the propensity score in observational studies for causal effects. **Biometrika,** 70 (1)**:** 41–55.

Rosenbaum, P. R. and Rubin, D. B. (1984) Reducing bias in observational studies using subclassification on the propensity score. **Journal of the American Statistical Association,** 79 (387)**:** 516–524.

Rosenbaum, P. R. and Rubin, D. B. (1985) The bias due to incomplete matching. **Biometrics,** 41 (1)**:** 103–116.

Rothman, K. J. (1986) "Multivariate analysis." **Modern epidemiology.** Boston, MA: Little, Brown and Company. pp.285–310.

Rothman, K. J. and Greenland, S. (1998) **Modern Epidemiology**. 2nd edition. Philadelphia: Lippincott-Raven.

Royal Infirmary of Edinburgh Renal Unit (2009) **How to start an ACE inhibitor** [online]. http://renux.dmed.ed.ac.uk/edren/Unitbits/ACEIstart.html [Accessed 3 July 2010].

Royston, P. (2005a) Multiple imputation of missing values: update. **Stata Journal,** 5 (2)**:** 1–14.

Royston, P. (2005b) Multiple imputation of missing values: Update of ice. **Stata Journal,** 5 (4)**:** 527–536.

Royston, P., Carlin, J. B. and White, I. R. (2009) Multiple imputation of missing values: New features for mim. **Stata Journal,** 9 (2)**:** 252–264.

Rubin, D. B. (1987) **Multiple imputation for non-response in surveys**. New York: John Wiley & Sons.

Rubin, D. B. (1997) Estimating causal effects from large data sets using propensity scores. **Annals of Internal Medicine,** 127 (8 Pt 2)**:** 757–763.

Rubin, D. B. (2007) The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. **Statistics in Medicine,** 26 (1)**:** 20–36.

Saito, M., Takada, M., Hirooka, K., et al. (2005) Serum concentration of potassium in chronic heart failure patients administered spironolactone plus furosemide and either enalapril maleate, losartan potassium or candesartan cilexetil. **Journal of Clinical Pharmacy and Therapeutics,** 30 (6)**:** 603–610.

Sauer, B. C., Kapoor, D., Nebeker, J. R., et al. (2006) Serum potassium monitoring in veterans using renin-angiotension-aldosterone inhibition. **Health Services Research & Development National Meeting. Arlington, Virginia, 2006 Feburary 16–17**.

Schafer, J. L. (1999) Multiple imputation: a primer. **Statistical Methods in Medical Research,** 8 (1)**:** 3–15.

Schafer, J. L. and Olsen, M. K. (1999) **Modeling and imputation of semicontinuous survey variables** [online]. The Methodology Center, Penn State University. www.fcsm.gov/99papers/shaffcsm.pdf [Accessed 20 May 2010].

Schneeweiss, S. (2006) Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. **Pharmacoepidemiology and Drug Safety,** 15 (5)**:** 291–303.

Schneeweiss, S. and Avorn, J. (2005) A review of uses of health care utilization databases for epidemiologic research on therapeutics. **Journal of Clinical Epidemiology,** 58 (4)**:** 323–337.

Schneeweiss, S., Hasford, J., Gottler, M., et al. (2002) Admissions caused by adverse drug events to internal medicine and emergency departments in hospitals: a longitudinal population-based study. **European Journal of Clinical Pharmacology,** 58 (4)**:** 285–291.

Scottish Intercollegiate Guidelines Network (2001) **Hypertension in Older People. SIGN Publication No. 49** [online]. Scottish Intercollegiate Guidelines Network, Edinburgh. www.sign.ac.uk/pdf/sign49.pdf [Accessed 25 July 2010].

Seedat, Y. K., Croasdale, M. A., Milne, F. J., et al. (2006) South African hypertension guideline 2006. **South African Medical Journal,** 96 (4 Pt 2)**:** 337–362.

Seeger, J. D., Walker, A. M., Williams, P. L., et al. (2003) A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. **American Journal of Cardiology,** 92 (12)**:** 1447–1451.

Sever, P., Beevers, G., Bulpitt, C., et al. (1993) Management guidelines in essential hypertension: report of the second working party of the British Hypertension Society. **BMJ,** 306 (6883)**:** 983–987.

Shah, B. R., Laupacis, A., Hux, J. E., et al. (2005) Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. **Journal of Clinical Epidemiology,** 58 (6)**:** 550–559.

Shah, N. R., Seger, A. C., Seger, D. L., et al. (2006) Improving acceptance of computerized prescribing alerts in ambulatory care. **Journal of the American Medical Informatics Association,** 13 (1)**:** 5–11.

Sharabi, Y., Illan, R., Kamari, Y., et al. (2002) Diuretic induced hyponatraemia in elderly hypertensive women. **Journal of Human Hypertension,** 16 (9)**:** 631–635.

SHEP Cooperative Research Group (1991) Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension. Final results of the Systolic Hypertension in the Elderly Program (SHEP). SHEP Cooperative Research Group. **JAMA,** 265 (24)**:** 3255–3264.

Sica, D. A. (2006) Antihypertensive therapy and its effects on potassium homeostasis. **Journal of Clinical Hypertension (Greenwich, Conn),** 8 (1)**:** 67–73.

Siegel, D., Hulley, S. B., Black, D. M., et al. (1992) Diuretics, serum and intracellular electrolyte levels, and ventricular arrhythmias in hypertensive men. **JAMA,** 267 (8)**:** 1083–1089.

Simon, S. R., Andrade, S. E., Ellis, J. L., et al. (2005) Baseline laboratory monitoring of cardiovascular medications in elderly health maintenance organization enrollees. **Journal of the American Geriatrics Society,** 53 (12)**:** 2165–2169.

Sin, D. D. and Tu, J. V. (2001) Inhaled corticosteroids and the risk of mortality and readmission in elderly patients with chronic obstructive pulmonary disease. **American Journal of Respiratory and Critical Care Medicine,** 164 (4)**:** 580–584.

Smellie, W. S., Forth, J., Coleman, J. J., et al. (2007) Best practice in primary care pathology: review 6. **Journal of Clinical Pathology,** 60 (3)**:** 225–234.

Smith, D. H., Raebel, M. A., Chan, K. A., et al. (2011) An economic evaluation of a laboratory monitoring program for renin-angiotensin system agents. **Medical Decision Making,** 31 (2)**:** 315–324.

Solaymani-Dodaran, M., Logan, R. F., West, J., et al. (2004) Risk of oesophageal cancer in Barrett's oesophagus and gastro-oesophageal reflux. **Gut,** 53 (8)**:** 1070–1074.

Soriano, J. B., Maier, W. C., Visick, G., et al. (2001) Validation of general practitioner-diagnosed COPD in the UK General Practice Research Database. **European Journal of Epidemiology,** 17 (12)**:** 1075–1080.

Souverein, P. C., Berard, A., van Staa, T. P., et al. (2004) Use of oral glucocorticoids and risk of cardiovascular and cerebrovascular disease in a population based case-control study. **Heart,** 90 (8)**:** 859–865.

Spital, A. (1999) Diuretic-induced hyponatremia. **American Journal of Nephrology,** 19 (4)**:** 447–452.

Stafford, R. S., Monti, V., Furberg, C. D., et al. (2006) Long-term and short-term changes in antihypertensive prescribing by office-based physicians in the United States. **Hypertension,** 48 (2)**:** 213–218.

Steele, A. W., Eisert, S., Witter, J., et al. (2005) The effect of automated alerts on provider ordering behavior in an outpatient setting. **PLoS Medicine,** 2 (9)**:** e255.

Sterne, J. A., White, I. R., Carlin, J. B., et al. (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. **BMJ,** 338**:** b2393.

Sterns, R. H. (1987) Severe symptomatic hyponatremia: treatment and outcome. A study of 64 cases. **Annals of Internal Medicine,** 107 (5)**:** 656–664.

Strom, B. L. (2005) "What is Pharmacoepidemiology?" In Strom, B. L. (ed.) **Pharmacoepidemiology.** 4th ed. Chichester: John Wiley & Sons Ltd. pp.3–15.

Stuart, E. A., Azur, M., Frangakis, C., et al. (2009) Multiple imputation with large data sets: a case study of the Children's Mental Health Initiative. **American Journal of Epidemiology,** 169 (9)**:** 1133–1139.

Stuart, E. A. and Rubin, D. B. (2008) "Best practices in quasi-experimental designs: why matched subject designs are superior to ANCOVA." In Osborne, J. W. (ed.) **Best Practices in Quantitative Methods.** Sage Publications, Inc.

Stürmer, T., Joshi, M., Glynn, R. J., et al. (2006a) A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. **Journal of Clinical Epidemiology,** 59 (5)**:** 437–447.

Stürmer, T., Rothman, K. J. and Glynn, R. J. (2006b) Insights into different results from different causal contrasts in the presence of effect-measure modification. **Pharmacoepidemiology and Drug Safety,** 15 (10)**:** 698–709.

Stürmer, T., Schneeweiss, S., Brookhart, M. A., et al. (2005) Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal antiinflammatory drugs and short-term mortality in the elderly. **American Journal of Epidemiology,** 161 (9)**:** 891–898.

Suissa, S. (2003) Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. **American Journal of Respiratory and Critical Care Medicine,** 168 (1)**:** 49–53.

Suissa, S. (2005) "Novel approaches to pharmacoepidemiology study design and statistical methods." In Strom, B. L. (ed.) **Pharmacoepidemiology.** 4th ed. Chichester: John Wiley & Sons Ltd. pp.811–829.

Suissa, S. (2007) Immortal time bias in observational studies of drug effects. **Pharmacoepidemiology and Drug Safety,** 16 (3)**:** 241–249.

Sylvestre, M. P., Huszti, E. and Hanley, J. A. (2006) Do OSCAR winners live longer than less successful peers? A reanalysis of the evidence. **Annals of Internal Medicine,** 145 (5)**:** 361–363; discussion 392.

Tannen, R. L., Weiner, M. G., Xie, D., et al. (2007a) Estrogen affects post-menopausal women differently than estrogen plus progestin replacement therapy. **Human Reproduction,** 22 (6)**:** 1769–1777.

Tannen, R. L., Weiner, M. G., Xie, D., et al. (2007b) A simulation using data from a primary care practice database closely replicated the women's health initiative trial. **Journal of Clinical Epidemiology,** 60 (7)**:** 686–695.

The ALLHAT Officers and Coordinators (2002) Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). **JAMA,** 288 (23)**:** 2981–2997.

Thomsen, L. A., Winterstein, A. G., Sondergaard, B., et al. (2007) Systematic review of the incidence and characteristics of preventable adverse drug events in ambulatory care. **Annals of Pharmacotherapy,** 41 (9)**:** 1411–1426.

Tleyjeh, I. M., Ghomrawi, H. M., Steckelberg, J. M., et al. (2010) Conclusion about the association between valve surgery and mortality in an infective endocarditis cohort changed after adjusting for survivor bias. **Journal of Clinical Epidemiology,** 63 (2)**:** 130–135.

Touw, D. J., Neef, C., Thomson, A. H., et al. (2005) Cost-effectiveness of therapeutic drug monitoring: a systematic review. **Therapeutic Drug Monitoring,** 27 (1)**:** 10–17.

Tragni, E., Filippi, A., Mazzaglia, G., et al. (2007) Monitoring statin safety in primary care. **Pharmacoepidemiology and Drug Safety,** 16 (6)**:** 652–657.

Uijtendaal, E. V., Zwart-van Rijkom, J. E., van Solinge, W. W., et al. (2011) Frequency of laboratory measurement and hyperkalaemia in hospitalised patients using serum potassium concentration increasing drugs. **European Journal of Clinical Pharmacology,** Epub ahead of print.

University of Michigan Health System (2009) **Essential hypertension: guidelines for clinical care** [online]. http://cme.med.umich.edu/pdf/guideline/htn.pdf [Accessed 2 April 2010].

van Buuren, S., Boshuizen, H. C. and Knook, D. L. (1999) Multiple imputation of missing blood pressure covariates in survival analysis. **Statistics in Medicine,** 18 (6)**:** 681–694.

van Buuren, S. and Oudshoorn, K. (1999) **Flexible multivariate imputation by MICE. TNO Report PG/VGZ/99.054**. Leiden: TNO Prevention and Health.

van der Sijs, H., Aarts, J., van Gelder, T., et al. (2008) Turning off frequently overridden drug alerts: limited opportunities for doing it safely. **Journal of the American Medical Informatics Association,** 15 (4)**:** 439–448.

van der Sijs, H., Aarts, J., Vulto, A., et al. (2006) Overriding of drug safety alerts in computerized physician order entry. **Journal of the American Medical Informatics Association,** 13 (2)**:** 138–147.

van Staa, T. P., Geusens, P., Pols, H. A., et al. (2005) A simple score for estimating the long-term risk of fracture in patients using oral glucocorticoids. **QJM,** 98 (3)**:** 191–198.

van Walraven, C., Davis, D., Forster, A. J., et al. (2004) Time-dependent bias was common in survival analyses published in leading clinical journals. **Journal of Clinical Epidemiology,** 57 (7)**:** 672–682.

van Wijk, B. L., Klungel, O. H., Heerdink, E. R., et al. (2006) Refill persistence with chronic medication assessed from a pharmacy database was influenced by method of calculation. **Journal of Clinical Epidemiology,** 59 (1)**:** 11–17.

Vinker, S., Kvint, I., Erez, R., et al. (2007) Effect of the characteristics of family physicians on their utilisation of laboratory tests. **British Journal of General Practice,** 57 (538)**:** 377–382.

Vittinghoff, E. and McCulloch, C. E. (2007) Relaxing the rule of ten events per variable in logistic and Cox regression. **American Journal of Epidemiology,** 165 (6)**:** 710–718.

Vrijens, B., Vincze, G., Kristanto, P., et al. (2008) Adherence to prescribed antihypertensive drug treatments: longitudinal study of electronically compiled dosing histories. **BMJ,** 336 (7653)**:** 1114–1117.

Waller, P. C. and Evans, S. J. (2003) A model for the future conduct of pharmacovigilance. **Pharmacoepidemiology and Drug Safety,** 12 (1)**:** 17–29.

Walley, T., Duggan, A. K., Haycox, A. R., et al. (2003) Treatment for newly diagnosed hypertension: patterns of prescribing and antihypertensive effectiveness in the UK. **Journal of the Royal Society of Medicine,** 96 (11)**:** 525–531.

Wandsworth NHS (2008) **Guidelines for the Initiation and Titration of Angiotensin-Converting Enzyme (ACE) Inhibitors** [online]. http://www.wandsworth-pct.nhs.uk/work/GPReferral/Referral/CVD/angina%20guidelines.asp [Accessed 3 March 2010].

Watson, D. J., Rhodes, T., Cai, B., et al. (2002) Lower risk of thromboembolic cardiovascular events with naproxen among patients with rheumatoid arthritis. **Archives of Internal Medicine,** 162 (10)**:** 1105–1110.

Weitzen, S., Lapane, K. L., Toledano, A. Y., et al. (2004) Principles for modeling propensity scores in medical research: a systematic literature review. **Pharmacoepidemiology and Drug Safety,** 13 (12)**:** 841–853.

Weitzen, S., Lapane, K. L., Toledano, A. Y., et al. (2005) Weaknesses of goodness-of-fit tests for evaluating propensity score models: the case of the omitted confounder. **Pharmacoepidemiology and Drug Safety,** 14 (4)**:** 227–238.

Wells, G. A., Shea, B., O'Connell, D., et al. (2010) **The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses** [online]. www.ohri.ca/programs/clinical_epidemiology/oxford.asp [Accessed 23 January 2010].

Wester, K., Jonsson, A. K., Spigset, O., et al. (2008) Incidence of fatal adverse drug reactions: a population based study. **British Journal of Clinical Pharmacology,** 65 (4)**:** 573–579.

White, I. R., Royston, P. and Wood, A. M. (2010) Multiple imputation using chained equations: Issues and guidance for practice. **Statistics in Medicine,** 30 (4)**:** 377–399.

White, I. R. and Thompson, S. G. (2005) Adjusting for partially missing baseline measurements in randomized trials. **Statistics in Medicine,** 24 (7)**:** 993–1007.

Whiting, P., Rutjes, A. W., Reitsma, J. B., et al. (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. **BMC Medical Research Methodology,** 3**:** 25.

Wiffen, P., Gill, M., Edwards, J., et al. (2002) Adverse drug reactions in hospital patients - a systematic review of the prospective and retrospective studies. **Bandolier Extra:** 1–16.

Williams, B., Poulter, N. R., Brown, M. J., et al. (2004) British Hypertension Society guidelines for hypertension management 2004 (BHS-IV): summary. **BMJ,** 328 (7440)**:** 634–640.

Wong, I. C., Ghaleb, M. A., Franklin, B. D., et al. (2004) Incidence and nature of dosing errors in paediatric medications: a systematic review. **Drug Safety,** 27 (9)**:** 661–670.

Xie, F., Petitti, D. B. and Chen, W. (2005) Prescribing patterns for antihypertensive drugs after the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial: report of experience in a health maintenance organization. **American Journal of Hypertension,** 18 (4 Pt 1)**:** 464–469.

Yanovitzky, I., Zanutto, E. and Hornik, R. (2005) Estimating causal effects of public health education campaigns using propensity score methods. **Evaluation and Program Planning,** 28**:** 209–220.

Yee, A., Majumdar, S. R., Simpson, S. H., et al. (2004) Statin use in Type 2 diabetes mellitus is associated with a delay in starting insulin. **Diabetic Medicine,** 21 (9)**:** 962–967.

Yusuf, S., Sleight, P. and Dagenais, G. (2000) **The HOPE (Heart Outcomes Prevention Evaluation) Study: a large randomized trial of the ACE inhibitor, ramipril, and Vitamin E in patients at high risk of cardiovascular events. FDA briefing document** [online]. www.fda.gov/ohrms/dockets/ac/00/backgrd/3612b1d.pdf [Accessed 4 April 2011].

Zanutto, E., Lu, B. and Hornik, R. (2005) Using propensity score subclassification for multiple treatment doses to evaluate a national antidrug media campaign. **Journal of Educational and Behavioral Statistics,** 30 (1)**:** 59–73.

Zhou, X. H., Eckert, G. J. and Tierney, W. M. (2001) Multiple imputation in public health research. **Statistics in Medicine,** 20 (9-10)**:** 1541–1549.

Zhou, Z. (2005) **Sensitivity of propensity score methods to the specifications. IZA discussion paper no. 1873**. Bonn, Germany: The Institute for the Study of Labor.

Zhou, Z., Rahme, E., Abrahamowicz, M., et al. (2005) Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. **American Journal of Epidemiology,** 162 (10)**:** 1016–1023.

Zoppi, M., Braunschweig, S., Kuenzi, U. P., et al. (2000) Incidence of lethal adverse drug reactions in the comprehensive hospital drug monitoring, a 20-year survey, 1974-1993, based on the data of Berne/St. Gallen. **European Journal of Clinical Pharmacology,** 56 (5)**:** 427–430.

Zuccalà, G., Pedone, C., Cocchi, A., et al. (2000) Older age and in-hospital development of hypokalemia from loop diuretics: results from a multicenter survey. GIFA Investigators. Multicenter Italian Pharmacoepidemiologic Study Group. **The Journals of Gerontology Series A, Biological Sciences and Medical Sciences,** 55 (4)**:** M232–238.

Appendices

Appendix 1 – Monitoring guidelines for patients treated with antihypertensive drugs

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
Treatment for hypertension			
SIGN (2001)	ACE inhibitors	Creatinine, potassium	- Serum creatinine and potassium must be checked within 1–2 weeks of commencing therapy
	AT-II receptor antagonists	Creatinine, potassium	- Serum potassium and creatinine should be checked within 1–2 weeks of commencing therapy
	Thiazide diuretic	Potassium	- Serum potassium levels should be checked within 4–6 weeks of starting low-dose therapy
Knight and Avorn (2001)	ACE inhibitors	Creatinine, potassium	- Serum potassium and creatinine levels should be checked within 1 week of initiation of therapy because this may prevent the development of renal insufficiency and hyperkalaemia
	Thiazide diuretics	Potassium	- Serum potassium should be checked within 1 week after initiation and at least yearly because of the risk for hypokalaemia due to diuretic therapy
Joint National Committee 7 (2003)	Antihypertensive therapy	Creatinine, potassium	- Serum potassium and creatinine tests are recommended before initiating therapy and should be monitored at least one to two times per year - Co-morbidities such as heart failure, associated diseases such as diabetes should influence the frequency of laboratory testing - With ACE inhibitors or AT-II receptor antagonist serum creatinine and potassium should be monitored 1–2 weeks following initiation or escalation in therapy in patients with renal transplantation

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
Palmer (2003)	ACE inhibitor or AT-II receptor antagonist	Potassium	- Serum potassium should be checked within 1–2 weeks of starting treatment in patients with chronic renal disease - If the potassium concentration increases to a value >5.6 mEq/l then another class of antihypertensive therapy will need to be utilized
Palmer (2004)	ACE inhibitor or AT-II receptor antagonist	Potassium	- Serum potassium concentration should be checked within one week after the drug has started - With each increase in the dose, the serum potassium concentration should be measured again one week later
British Hypertension Society (2004)	Antihypertensive drugs in general	Creatinine Electrolytes	- Creatinine monitoring is described as a routine investigation - Potassium monitoring is described as a routine investigation
French Haute Autorité de Santé (2005)	Antihypertensive drugs	Biochemical tests	- Serum creatinine, sodium, and potassium should be measured before initiation of treatment - If an ACE inhibitor or AT-II receptor antagonist is prescribed, tests for serum potassium and creatinine should be done within 7–15 days of starting treatment - If serum creatinine rises by more than 20–30% while on an ACE inhibitor or AT-II receptor antagonist, treatment should be discontinued and the patient referred to a specialist - Serum creatinine, sodium, and potassium should then be confirmed at least twice a year and in the event of intercurrent disease - Serum creatinine should be specifically monitored in elderly patients treated with diuretics, ACE inhibitors and AT-II receptor antagonists, and/or prescribed in combination with potentially nephrotoxic drugs

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
South African Medical Association (2006)	Antihypertensive drugs	Creatinine and potassium	- Tests of creatinine and potassium are described as routine investigations and should be undertaken yearly if normal
ESH/ESC (2007)	Antihypertensive drugs	Creatinine and potassium	- Tests of creatinine and potassium are described as routine investigations *Additional information:* - The combination of ACE inhibitors and AT-II receptor antagonists in combination or of high doses of angiotensin receptor antagonists may be used if careful attention is paid to possible rises in serum creatinine and potassium.
Smellie (2007)	ACE inhibitors; AT-II receptor antagonists	Creatinine, electrolytes	- Before initiating treatment - 1 week after starting treatment or any subsequent dose increase - At 4 and 10 days after starting treatment or increase in dose in patients at higher risk of developing hyperkalaemia or deteriorating renal function (e.g., peripheral vascular disease, diabetes mellitus, pre-existing renal impairment and older patients) Consider seeking further advice if a patient has: - Renal impairment (serum creatinine >200 mmol/l or eGFR <30 ml/min) or confirmed/suspected renovascular disease before initiating ACE inhibitor/AT-II receptor antagonist - Marked creatinine rise (>30%) with large fall in blood pressure after starting ACE inhibitor or AT-II receptor antagonist may suggest renovascular disease that should be investigated
	Thiazide or loop diuretic	Creatinine, electrolytes	- Within 4–6 weeks of starting low-dose thiazide diuretic treatment or loop diuretic treatment - thereafter, in all patients every 6–12 months - or if a person's clinical condition changes or a potentially interacting drug is added

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
	Spironolactone or potassium-sparing diuretics	Creatinine, electrolytes	- before initiation of treatment (it should not be initiated if the potassium values >5 mmol/l) - after 5–7 days with dose titration if required - every 5–7 days until the potassium values are stable - 1–2 times/year up to every 4–8 weeks during chronic treatment, depending on risk factors (older patients, renal or cardiac dysfunction) - If potassium rises to >6 mmol/l, spironolactone or potassium-sparing diuretics should be stopped and specialist advice sought
Northern Ireland Department of Health, Social Services and Public Safety (2007)	ACE inhibitors, AT-II receptor antagonists	Creatinine and electrolytes	- Before starting treatment, check electrolytes and renal function - Recheck electrolytes and creatinine within 2 weeks of starting or increasing the dose - A rise in serum creatinine concentration of more than 30% after initiation of therapy or a dose increase should be followed by further measurements within 2 weeks - Every 12 months Notes - If serum potassium levels rise to above 6.0 mmol/l stop nephrotoxic drugs, potassium-sparing diuretics
British Columbia Ministry of Health Services (2008)	ACE inhibitors, AT-II receptor antagonists	Creatinine and electrolytes	- Before starting treatment, check creatinine, sodium and potassium - If combining ACE inhibitor and AT-II receptor antagonist, monitor for hyperkalaemia and worsening renal function

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
Australian Heart Foundation (2008)	Antihypertensive drugs	Biochemical tests	- Initial tests of sodium, potassium and creatinine should be carried out. - Monitor ACE inhibitors for hyperkalaemia, hyponatraemia - Monitor AT-II receptor antagonists for hyperkalaemia - Monitor thiazide diuretics for hyperkalaemia, hyponatraemia Notes: An initial rise in serum creatinine is commonly observed after initiation of ACE inhibitors or AT-II receptor antagonists. An increase of 30% or less is acceptable. If creatinine increases by more than 30% from baseline, consider possible contributory factors (e.g. hypovolaemia, renal artery stenosis, NSAIDs). If none present, consider ceasing treatment. Do not commence these agents if serum potassium is > 5.0 mmol/L.
University of Michigan Health System (2009)	Antihypertensive therapy	Creatinine and potassium	- Consider tests of creatinine and potassium before therapy is initiated. - Serum potassium and creatinine should be monitored at least 1–2 times/year.
Michigan Quality Improvement Consortium (2009)	Antihypertensive therapy	Creatinine and potassium	- Measure potassium and creatinine prior to initiating therapy
Royal Infirmary of Edinburgh Renal Unit (2009)	ACE inhibitor	Creatinine	- Before starting treatment - Check at 4 and 10 days after initiating treatment in patients with peripheral vascular disease, diabetes mellitus, old age, and pre-existing renal impairment - Check at 7 days after initiating treatment in patients without the risk factors described above

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
		Potassium	- Before starting treatment - Check at 4 and 10 days after initiating treatment in patients with peripheral vascular disease, diabetes mellitus, old age, and pre-existing renal impairment - Check at 7 days after initiating treatment in patients without the risk factors described above *Specific guidance:* - If 5.0–5.5 mmol/l, recheck in 7 days; 5.6–6.0 mmol/l stop ACE inhibitor and check in 7 days; 6.1–6.5 mmol/l stop ACE inhibitor and check immediately; >6.5 mmol/l stop ACE inhibitor and check urgently
		Sodium and urea	- Before starting treatment
Canadian Hypertension Education Program (2009)	Antihypertensive drugs	Blood chemistry (potassium, sodium, and creatinine)	- Routine laboratory tests should be performed for the investigation of all patients with hypertension - Monitor potassium and renal function of combining an ACE inhibitor with an AT-II receptor antagonists - Patients with non-diabetic chronic kidney disease place on an ACE inhibitor of an AT-II receptor antagonist should have their serum creatinine and potassium carefully monitored
Japanese Society of Hypertension (2009)	Antihypertensive drugs	Creatinine and potassium	- Attention to hyperkalaemia is necessary while using AT-II receptor antagonists with a potassium-sparing diuretic - Caution is necessary in patients with renal dysfunction because hyperkalaemia is most likely to occur 1–2 days after commencing treatment with ACE inhibitors. ACE inhibitors should not be used in patients suggested to have bilateral renovascular hypertension or renovascular hypertension with a functionally solitary kidney, as they may cause renal failure. If they are used, monitoring of the serum creatinine and potassium levels is necessary.
Institute for Clinical Systems Improvement (2010)	Antihypertensive drugs	Creatinine and electrolytes	- Initial laboratory tests prior to the initiation of therapy include: sodium, potassium, and creatinine

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
NCQA (2010)	ACE inhibitors, AT-II receptor antagonists, diuretics	Creatinine and potassium	- Serum creatinine and potassium should be measured annually
In chronic kidney disease			
National Kidney Foundation (2002)	ACE inhibitors, AT-II receptor antagonists, diuretics	Potassium	- At initiation and increase in dose of ACE inhibitor, AT-II receptor antagonist or diuretic, the levels of serum potassium should be measured to establish a "baseline" or "new baseline." Follow-up measurements should be undertaken as follows: *Time* — *Baseline serum potassium* Every 4–12 weeks — ≤4.5 mEq/L Every 2–4weeks — 4.6–5.0 mEq/L ≤2 weeks — >5.0 mEqL - After blood pressure is at goal and dose is stable , follow-up measurements should be made as follows: *Time* — *Baseline serum potassium* 6–1months — ≤4.5 mEq/L 3–6 months — 4.6–5.0 mEq/L 1–3 months — >5.0 mEq/L - If hyperkalaemia develops, reduce the dose of ACE inhibitor or AT-II receptor antagonist by 50% and reassess the serum potassium every 5 to 7 days until serum potassium has returned to baseline. If serum potassium does not return to baseline within 2 to 4 weeks, discontinue the ACE inhibitor or AT-II receptor antagonist and select an alternate antihypertensive agent.

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
VA/DoD (2007)	ACE inhibitors, AT-II receptor antagonists	Creatinine and potassium	- Creatinine and potassium levels should be monitored one to two weeks after initiation or after a change in dose of ACEI or AT-II receptor antagonist therapy and periodically to maintain a normal range. - After initiating an ACEI or AT-II receptor antagonist, it is recommended that the patient's potassium be checked within 2 weeks if baseline was > 5.0 mEq/L, at 2 to 4 weeks if baseline potassium was 4.5–5.0 mEq/L, and at 4 to 12 weeks if baseline was < 4.5 mEq/L - In most patients, an ACEI or AT-II receptor antagonist should be continued unless: a. There is an acute GFR decline of > 30 percent within the first two weeks after initiation. b. Serum potassium is > 6 mEq/L, despite appropriate treatment.

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
NICE (2008)	ACE inhibitors, AT-II receptor antagonists	Creatinine and potassium	- Measure serum potassium and creatinine concentrations before starting ACE inhibitor/AT-II receptor antagonist therapy. - Repeat these measurements between 1 and 2 weeks after starting ACE inhibitor/ARB therapy and after each dose increase. *Additional guidance:* - ACE inhibitor/AT-II receptor antagonist therapy should not normally be started if the pretreatment serum potassium concentration is significantly above the normal reference range (typically more than 5.0 mmol/litre). - Concurrent prescription of drugs known to promote hyperkalaemia is not a contraindication to the use of ACE inhibitors/AT-II receptor antagonists, but be aware that more frequent monitoring of serum potassium concentration may be required. - Stop ACE inhibitor/AT-II receptor antagonist therapy if the serum potassium concentration rises to 6.0 mmol/litre or more and other drugs known to promote hyperkalaemia have been discontinued.

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
Treatment in heart failure			
McMurray (2005)	ACE inhibitors and AT-II receptor antagonists	Creatinine, potassium	- Monitor creatinine and potassium 1–2 weeks after initiation and 1–2 weeks after final dose titration - Some rise in creatinine and potassium is to be expected after initiation of an ACE inhibitor; if an increase is small and asymptomatic no action is necessary - An increase in creatinine of up to 50% above baseline, or 266 mmol/L (3 mg/dL), whichever is the smaller, is acceptable - An increase in potassium to ≤5.5 mmol/L is acceptable - If creatinine or potassium do rise excessively consider stopping concomitant nephrotoxic drugs (e.g., NSAIDs), other potassium supplements/retaining agents (triamterene, amiloride, spironolactone/eplerenone) and, if no signs of congestion, reducing the dose of diuretic - If greater rises in creatinine or potassium than those outlined above persist despite adjustment of concomitant medications, the dose of the ACE inhibitor should be halved and blood chemistry rechecked within 1–2 weeks; if there is still an unsatisfactory response specialist advice should be sought - If potassium rises to >5.5 mmol/L or creatinine increases by >100% or to above 310 ,mol/L (3.5 mg/dL) the ACE inhibitor should be stopped and specialist advice sought - Blood chemistry should be monitored frequently and serially until potassium and creatinine have plateaued
North East Essex Medicines Management Committee (2007)	ACE inhibitors	Urea, creatinine, electrolytes	- Check after one week. - Review after one month - Review at 6 months (or before if there is a dose change or patients become unwell) - At annual review

Guideline	Drug class	Monitoring test(s)	Monitoring recommendations
Wandsworth Primary Care Trust NHS (2008)	ACE inhibitors	Creatinine and electrolytes	- Check baseline renal function (creatinine) and electrolytes - Re-check renal function and electrolytes at week one, two, and four after initiation - Once titrated, continue monitoring creatinine and electrolytes every 6 months *Notes* - Some rise in urea, creatinine and potassium is to be expected after initiation of an ACE inhibitor (an increase in creatinine of up to 50% above baseline or up to 200 µmol/l and an increase in potassium to ≤5.9 mmol/l is acceptable) - If potassium rises ≥6.0 mmol/l or creatinine rises by >100% or to 350 µmol/l, stop ACE inhibitor and seek advice
Eccles (1998)	ACE inhibitors	Creatinine and potassium	- Serum creatinine and potassium should be tested before initiation of treatment, one week after the start of treatment, and again one week after each significant increase in dosage
Lloyd and Mauro (2000)	Spironolactone	Creatinine and potassium	- Serum creatinine and potassium concentrations should be monitored after 7 days of treatment and the frequently (weekly–monthly) for the first few months, and routinely (every 3–6 months) thereafter.

Appendix 2 – Monitoring recommendations from SPCs for antihypertensive drugs

Drug class	Drug	Monitoring recommendations
Alpha blockers	Doxazosin	None
	Indoramin	None
	Prazosin	None
	Terazosin	None
ACE inhibitors	Captopril	- Careful titration and monitoring of renal function should be carried out in patients with bilateral renal artery stenosis; - Routine monitoring of potassium and creatinine is part of normal medical practice in patients with renal impairment; - Regular monitoring of serum potassium is recommended with concomitant treatment with potassium-sparing diuretics.
	Cilazapril	- Renal function in patients with renal impairment should be monitored during the first weeks; - Routine monitoring of potassium and creatinine is part of normal medical practice for these patients; - Potassium-sparing diuretics should be used with caution and serum potassium and renal function should be monitored frequently; - Patients treated with concomitant NSAID should have monitoring of their renal function after initiation of concomitant therapy, and periodically thereafter.
	Enalapril	- Renal function and serum potassium should be monitored in patients with hypertension. - Renal function should be monitored closely both before and after starting treatment in patients with heart failure/asymptomatic left ventricular dysfunction. - Routine monitoring of potassium and creatinine are part of normal medical practice for patients with renal function impairment. - Concomitant use of potassium-sparing diuretics and potassium supplements should be used with caution and with frequent monitoring of serum potassium.
	Fosinopril	- Concomitant use of potassium-sparing diuretics should be used with caution and the patient's serum potassium should be monitored frequently. - Renal function should be assessed prior to initiation of therapy and during treatment where appropriate.
	Imidapril	- Close monitoring of renal function during therapy should be performed as deemed appropriate in patients with renal insufficiency. - It is recommended that the renal function be monitored during the first weeks of therapy. - Monitoring of renal function is recommended in patients with renovascular hypertension. - Regular monitoring of serum potassium is recommended with the use of concomitant potassium supplements or potassium-sparing diuretics. - Consideration should be given to monitoring renal function after initiation of concomitant therapy with NSAIDs, and periodically thereafter.

Drug class	Drug	Monitoring recommendations
	Lisinopril	- Renal function and serum potassium should be monitored in patients treated concomitantly with diuretics. - In patients with renal failure, renal function should be monitored in the first weeks of therapy. - Routine monitoring of potassium and creatinine is part of normal medical practice for patients with renal failure. - Regular monitoring of serum potassium is recommended with concomitant use of potassium-sparing diuretics, potassium supplements or potassium-containing salt substitutes.
	Moexipril	- Especially at the beginning of the ACE inhibitor therapy the blood pressure and the respective laboratory values must be monitored carefully in patients with: impaired renal function (creatinine clearance 40-60 ml/min); renal hypertension; cardiac failure; salt and/or fluid volume depletion; or age of more than 65 years. - In hypertensive patients with renal artery stenosis in a solitary kidney or bilateral renal artery stenosis, renal function should be monitored during the first few weeks of therapy. - Concomitant use with potassium-sparing diuretics (e.g. spironolactone, triamterene, amiloride) potassium supplements or potassium-containing salt substitutes should be given with caution and with frequent monitoring of serum potassium.
	Perindopril	- Renal function and serum potassium should be monitored. - Patients at high risk of symptomatic hypotension e.g. patients with salt depletion with or without hyponatraemia, patients with hypovolaemia or patients who have been receiving vigorous diuretic therapy should have these conditions corrected, if possible, prior to therapy with perindopril. - Blood pressure, renal function and serum potassium should be monitored closely, both before and during treatment with perindopril. - Routine monitoring of potassium and creatinine are part of normal medical practice for patients with renal impairment. - Frequent monitoring of serum potassium is recommended when the following drugs are used concomitantly: potassium-sparing diuretics (e.g. spironolactone, eplerenone, triamterene, or amiloride), potassium supplements or potassium-containing salt substitutes; or those patients taking other drugs associated with increases in serum potassium (e.g. heparin). - Use of ACE-inhibitors and NSAIDs requires monitoring of renal function and serum potassium after initiation of concomitant therapy, and periodically thereafter.
	Quinapril	- In patients with renal insufficiency, monitoring of renal function during therapy should be performed as deemed appropriate; although in the majority renal function will not alter or may improve. - Specifically, patients with a creatinine clearance of <40 ml/min require a lower initial dosage of quinapril and their renal function should be closely monitored. - In hypertensive patients with unilateral or bilateral renal artery stenosis, renal function should be monitored during the first few weeks of therapy. - Concomitant treatments with potassium-sparing diuretics, potassium supplements or potassium salts should be used with caution and with appropriate monitoring of serum potassium.

Drug class	Drug	Monitoring recommendations
	Ramipril	- Patients concomitantly treated with diuretics should have their renal function and serum potassium monitored. - Renal function should be assessed before and during treatment and dosage adjusted especially in the initial weeks of treatment. - Particularly careful monitoring is required in patients with renal impairment. - Regular monitoring of serum potassium is recommended in patients at risk for development of hyperkalaemia include those with renal insufficiency, age (> 70 years), uncontrolled diabetes mellitus, or those using potassium salts, potassium retaining diuretics and other plasma potassium increasing active substances, or conditions such as dehydration, acute cardiac decompensation, metabolic acidosis.
	Trandolapril	- Patients with severe renal insufficiency may require reduced doses of trandalopril; their renal function should be closely monitored. - Concomitant use with potassium-sparing diuretics (spironolactone, amiloride, triamterene) or potassium supplements requires regular monitoring of serum potassium.
AT-II receptor antagonists	Candesartan	- Evaluation of patients with HF should always comprise assessment of renal function, especially in elderly patients 75 years or older, including monitoring of serum creatinine and potassium. - When used in hypertensive patients with renal impairment, periodic monitoring of serum potassium and creatinine levels is recommended. - During dose titration, monitoring of serum creatinine and potassium is recommended. - Regular monitoring should be undertaken with used in combination with an ACE inhibitor in HF. - Concomitant use with potassium-sparing diuretics, potassium supplements requires monitoring of potassium as appropriate. - Concomitant treatment with a NSAID may increase risk of worsening renal function and consideration should be given to monitoring renal function after initiation of concomitant treatment. - In patients with renal impairment, periodic monitoring of serum potassium and creatinine levels is recommended.
	Eprosartan	- Regular monitoring for serum potassium levels is recommended with concomitant use of potassium supplements or potassium-sparing diuretics. - Monitoring of renal function after the initiation of concomitant NSAID, and periodically thereafter, is recommended. - When eprosartan is to be used in patients with renal impairment, renal function should be assessed before starting treatment and at intervals during the course of therapy. If worsening of renal function is observed during therapy, treatment with eprosartan should be reassessed.

Drug class	Drug	Monitoring recommendations
	Irbesartan	- When irbesartan is used in patients with impaired renal function, a periodic monitoring of potassium and creatinine serum levels is recommended. - Close monitoring of serum potassium in patients at risk of hyperkalaemia (renal impairment, overt proteinuria due to diabetic renal disease, and/or heart failure) is recommended. - Patients should be adequately hydrated and consideration should be given to monitoring renal function after initiation of concomitant therapy with NSAIDs, and periodically thereafter.
	Losartan	- Plasma concentrations of potassium as well as creatinine clearance values should be closely monitored, especially patients with heart failure and a creatinine clearance between 30–50 ml/ min should be closely monitored. - Renal function should be regularly monitored during treatment with losartan as it may deteriorate. This applies particularly when losartan is given in the presence of other conditions (fever, dehydration) likely to impair renal function. - Patients should be adequately hydrated and consideration should be given to monitoring renal function after initiation of concomitant therapy with NSAIDs, and periodically thereafter.
	Olmesartan	- When olmesartan is used in patients with impaired renal function, periodic monitoring of serum potassium and creatinine levels is recommended. - Close-monitoring of serum potassium in patients at risk of hyperkalaemia is recommended: diabetes, renal impairment, age (> 70 years); combination with one or more other medicinal products that affect the renin-angiotensin-aldosterone system and/or potassium supplements. Some medicinal products or therapeutic class of medicinal products may provoke a hyperkalaemia: salt substitutes containing potassium, potassium-sparing diuretics, ACE inhibitors, angiotensin II receptors antagonists, non steroidal anti-inflammatory drugs (including selective COX-2 inhibitors), heparin, immunosuppressor as ciclosporin or tacrolimus, trimethoprim; or intercurrent events, in particular dehydration, acute cardiac decompensation, metabolic acidosis, worsening of renal function, sudden worsening of the renal condition (e.g. infectious diseases), cellular lysis (e.g. acute limb ischemia, rhabdomyolysis, extended trauma). - Close monitoring of renal function is advised in hepatically-impaired patients who are already receiving diuretics and/or other antihypertensive agents.

Drug class	Drug	Monitoring recommendations
	Telmisartan	- When telmisartan is used in patients with impaired renal function, periodic monitoring of potassium and creatinine serum levels is recommended. - Dual blockade of the renin-angiotensin-aldosterone system (e.g. by adding an ACE-inhibitor to an angiotensin II receptor antagonist) is not recommended in patients with already controlled blood pressure and should be limited to individually defined cases with close monitoring of renal function. - Close monitoring of serum potassium in at patients at risk for hyperkalaemia: diabetes mellitus, renal impairment, age (>70 years); combination with one or more other medicinal products that affect the renin-angiotensin-aldosterone system and/or potassium supplements. Medicinal products or therapeutic classes of medicinal products that may provoke hyperkalaemia are salt substitutes containing potassium, potassium-sparing diuretics, ACE inhibitors, angiotensin II receptor antagonists, non steroidal anti-inflammatory medicinal products (NSAIDs, including selective COX-2 inhibitors), heparin, immunosuppressives (cyclosporin or tacrolimus), and trimethoprim; or intercurrent events, in particular dehydration, acute cardiac decompensation, metabolic acidosis, worsening of renal function, sudden worsening of the renal condition (e.g. infectious diseases), cellular lysis (e.g. acute limb ischemia, rhabdomyolysis, extend trauma). - Frequent monitoring of serum potassium is necessary with concomitant use of potassium-sparing diuretics e.g. spironolactone, eplerenone, triamterene, or amiloride, potassium supplements, or potassium-containing salt substitutes may lead to a significant increase in serum potassium. - Concomitant use of NSAIDs requires that patients should be adequately hydrated and consideration should be given to monitoring of renal function after initiation of concomitant therapy and periodically thereafter.
	Valsartan	- Other agents that affect the renin-angiotensin system may increase blood urea and serum creatinine in patients with unilateral renal artery stenosis, therefore monitoring of renal function is recommended when patients are treated with valsartan. - If a medicinal product that affects potassium levels (e.g. potassium-sparing diuretics, potassium supplements, salt substitutes containing potassium and other substances that may increase potassium levels) is considered necessary in combination with valsartan, monitoring of potassium plasma levels is advised. - Monitoring of renal function and serum potassium at the beginning of the treatment is recommended with concomitant treatment with NSAIDs, as well as adequate hydration of the patient.
Beta blockers	Acebutolol	No mention
	Atenolol	No mention
	Betaxolol	No mention
	Bisoprolol	No mention
	Carvedilol	- Renal function should be monitored during up-titration in patients with CHF and the drug discontinued or dosage reduced if worsening of renal function occurs.

Drug class	Drug	Monitoring recommendations
	Celiprolol	- Close monitoring of elderly patients should be exercised, as renal and hepatic functions may be decrease in this population.
	Metoprolol	No mention
	Nadolol	No mention
	Nebivolol	No mention
	Oxprenolol	No mention
	Pindolol	No mention
	Propranolol	No mention
	Sotalol	- Potassium levels should be monitored and corrected appropriately during concomitant administration with other potassium-depleting drugs: Amphotericin B (IV route), corticosteroids (systemic administration), and some laxatives.
	Timolol	No mention
Calcium channel blockers	Amlodipine	No mention
	Diltiazem	No mention
	Felodipine	No mention
	Isradipine	No mention
	Lacidipine	No mention
	Lercanidipine	No mention
	Nicardipine	No mention
	Nifedipine	No mention
	Nimodipine	No mention
	Verapamil	No mention
Loop diuretics	Bumetanide	- Regular checks of serum electrolytes, in particular sodium, potassium, chloride, and bicarbonate, should be performed.
	Furosemide	- Regular monitoring of creatinine and electrolyte balance is recommended. - Frequent checks of the serum potassium level are necessary in patients with impaired renal function.
	Torasemide	- On long-term treatment with torasemide, regular monitoring of the electrolyte balance, glucose, uric acid, creatinine and lipids in the blood, is recommended.

Drug class	Drug	Monitoring recommendations
Potassium-sparing diuretics and aldosterone antagonists	Eplerenone	- Renal function and serum potassium should be monitored in patients with hypertension. - Renal function should be monitored closely both before and after starting treatment in patients with heart failure/asymptomatic left ventricular dysfunction. - Routine monitoring of potassium and creatinine are part of normal medical practice for patients with renal function impairment. - Concomitant use of potassium-sparing diuretics and potassium supplements should be used with caution and with frequent monitoring of serum potassium.
Potassium-sparing diuretics with other diuretics	Amiloride	- Patients with increased in blood urea over 10 mmol/l, serum creatinine over 130 micromol/l, or with diabetes mellitus, should not received amiloride without careful, frequent monitoring of serum electrolytes and blood urea levels. - When administered concomitantly with an ACE inhibitor, ARB, trilostrane, ciclosporin or tacrolimus, frequent monitoring of serum potassium is recommended. - When used concomitantly with NSAIDs, renal function and serum potassium levels should be carefully monitored. - When given with other diuretics, careful monitoring of serum electrolyte and blood urea levels should be carried out.
	Spironolactone	- Fluid and electrolyte status should be regularly monitored particularly in the elderly, in those with significant renal and hepatic impairment. - Hyperkalaemia may occur in patients with impaired renal function or excessive potassium intake and can cause cardiac irregularities which may be fatal. Should hyperkalaemia develop, spironolactone should be discontinued, and if necessary, active measures taken to reduce the serum potassium to normal. - Concurrent administration of angiotensin-II receptor antagonists, e.g. valsartan, losartan, and spironolactone may result in an increase in serum potassium levels. If concurrent use is necessary, monitor serum potassium levels. - Avoid concurrent use of spironolactone and ciclosporin. If concurrent therapy is necessary, monitor serum potassium levels for persistent elevations in patients. - Co-administration of spironolactone with fludrocortisone may result in a paradoxical dose-related increase in urinary potassium excretion. If concomitant administration is necessary, closely monitor serum potassium levels. - Potassium supplements are contraindicated except in cases of initial potassium depletion. If potassium supplementation is considered essential, serum electrolytes should be monitored.
	Triamterene	- It is advisable to monitor blood urea, serum potassium levels and electrolytes periodically. This is important in the elderly, those with renal impairment and those receiving concomitant treatment with NSAIDs. - It is advisable to monitor blood urea and serum potassium levels periodically in patients receiving concomitant treatment with NSAIDs. - Monitor serum potassium during first cycle of drospirenone.

Drug class	Drug	Monitoring recommendations
Thiazides and related diuretics	Bendro-flumethiazide	- Renal function should be continuously monitored during thiazide therapy. Serum electrolytes should be checked for abnormalities, particularly hypokalaemia. Elderly patients and patients who are on long term treatment need regular blood tests to monitor electrolyte levels. - Bendroflumethiazide may enhance the nephrotoxicity of NSAIDs. Indometacin and ketorolac antagonise the diuretic effect of bendroflumethiazide, this occurs to a lesser extent with ibuprofen, piroxicam and naproxen. The effects of concurrent use should be monitored and the dose of bendroflumethiazide modified if necessary. - There is an increased risk of hypokalaemia and a decrease in diuretic activity when carbenoxolone and bendroflumethiazide are taken together. Patients should be monitored and given potassium supplements when required.
	Chlortalidone	- Periodic serum electrolyte determinations should be carried out. - Monitoring of serum electrolytes is particularly indicated in the elderly, in patients with ascites due to liver cirrhosis, and in patients with oedema due to nephrotic syndrome.
	Indapamide	- Plasma sodium must be measured before starting treatment, then at regular intervals subsequently. - Any diuretic treatment may cause hyponatraemia. The fall in plasma sodium may be asymptomatic initially and regular monitoring is therefore essential, and should be even more frequent in the elderly and cirrhotic patients. - Plasma potassium should be first measured during the first week following the start of treatment. - More frequent monitoring should be targetted at patients at high risk of hypokalaemia: elderly, malnourished and/or polymedicated, cirrhotic patients with oedema and ascites, coronary artery disease and cardiac failure patients, individuals with a long QT interval. Monitor renal function at the start of treatment with NSAIDs, baclofen, or ACE inhibitors used concomitantly. Monitor plasma potassium if used concomitantly with amphotericin B, gluco- and mineralocorticoids, tetracosactide, stimulant laxatives, digitalis, or potassium-sparing diuretics.
	Metolazone	- Fluid and electrolyte balance should be carefully monitored during therapy especially if metolazone is used concurrently with other diuretics. - In particular, metolazone may potentiate the diuresis produced by furosemide and, if the two agents are used concurrently, patients should be carefully monitored. - Prolonged therapy with metolazone may result in hypokalaemia. Serum potassium levels should be determined at regular intervals and, if necessary, potassium supplementation should be instituted.
	Xipamide	No mention

Appendix 3 – Medline and Embase search strategies used to identify monitoring studies

Medline search strategy

Search term
1. *Angiotensin-Converting Enzyme Inhibitors/
2. *Angiotensin II Type 1 Receptor Blockers/
3. *Calcium Channel Blockers/
4. *Thiazides/
5. *Diuretics/
6. *Adrenergic alpha-Antagonists/
7. *Adrenergic beta-Antagonists/
8. *Aldosterone Antagonists/
9. *Antihypertensive Agents/
10. angiotensin converting enzyme inhibitors.ab,ti.
11. angiotensin receptor blockers.ab,ti.
12. angiotensin II receptor antagonists.ab,ti.
13. calcium channel blockers.ab,ti.
14. thiazide diuretics.ab,ti.
15. potassium sparing diuretics.ti,ab.
16. loop diuretics.ti,ab.
17. alpha adrenoceptor blocking drugs.ti,ab.
18. beta adrenoceptor blocking drugs.ti,ab.
19. alpha blockers.ti,ab.
20. beta blockers.ti,ab.
21. aldosterone antagonists.ti,ab.
22. potassium sparing diuretics.ti,ab.
23. antihypertensive drug$.ti,ab.
24. OR/1–23
25. *Drug Monitoring/
26. *Monitoring, Physiologic/
27. (laboratory adj5 monitoring).ti,ab.
28. (biochemical adj5 monitoring).ti,ab.
29. (sodium adj5 concentration).ti,ab. OR (potassium adj5 concentration).ti,ab. OR (creatinine adj5 concentration).ti,ab. OR (urea adj5 concentration).ti,ab.
30. (sodium adj5 test).ti,ab. OR (potassium adj5 test).ti,ab. OR (creatinine adj5 test).ti,ab. OR (urea adj5 test).ti,ab.
31. (monitoring adj100 creatinine).ti,ab. OR (monitoring adj100 electrolyte).ti,ab. OR (monitoring adj100 urea).ti,ab. OR (monitoring adj100 potassium).ti,ab. OR (monitoring adj100 sodium).ti,ab. OR (monitoring adj100 renal function).ti,ab.
32. OR/25–31
33. 24 AND 32

Embase search strategy

Search term
1. *beta adrenergic receptor blocking agent/
2. *angiotensin 2 receptor antagonist/ OR *angiotensin 1 receptor antagonist/ OR *angiotensin receptor antagonist/
3. *perindopril/ OR *captopril/ OR *enalapril/ OR *lisinopril/ OR *cilazapril/
4. *calcium channel blocking agent/
5. *diuretic agent/
6. *aldosterone antagonist/
7. *alpha adrenergic receptor blocking/ OR *alpha adrenergic receptor blocking agent/ OR *alpha 2 adrenergic receptor blocking agent/ OR *alpha 1 adrenergic receptor blocking agent/
8. *loop diuretic agent/
9. *potassium sparing diuretic agent/
10. *antihypertensive agent/
11. angiotensin converting enzyme inhibitors.ab,ti.
12. angiotensin receptor blockers.ab,ti.
13. angiotensin II receptor antagonists.ab,ti.
14. calcium channel blockers.ab,ti.
15. thiazide diuretics.ab,ti.
16. potassium sparing diuretics.ti,ab.
17. loop diuretics.ti,ab.
18. alpha adrenoceptor blocking drugs.ti,ab.
19. beta adrenoceptor blocking drugs.ti,ab.
20. alpha blockers.ti,ab.
21. beta blockers.ti,ab.
22. aldosterone antagonists.ti,ab.
23. potassium sparing diuretics.ti,ab.
24. antihypertensive drug$.ti,ab.
25. OR/1–24
26. *patient monitoring/
27. *drug monitoring/
28. *ambulatory monitoring/
29. (laboratory adj5 monitoring).ti,ab.
30. (biochemical adj5 monitoring).ti,ab.
31. (sodium adj5 concentration).ti,ab. OR (potassium adj5 concentration).ti,ab. OR creatinine adj5 concentration).ti,ab. OR (urea adj5 concentration).ti,ab.
32. (sodium adj5 test).ti,ab. OR (potassium adj5 test).ti,ab. OR (creatinine adj5 test).ti,ab. OR (urea adj5 test).ti,ab.
33. (monitoring adj100 creatinine).ti,ab. OR (monitoring adj100 electrolyte).ti,ab. OR (monitoring adj100 urea).ti,ab. OR (monitoring adj100 potassium).ti,ab. OR (monitoring adj100 sodium).ti,ab. OR (monitoring adj100 renal function).ti,ab.
34. OR/26–33
35. 25 AND 34

Appendix 4 – Data extraction form for systematic review of monitoring studies

Study name	
Year	
Country	
Patient characteristics	Total number: Setting: Age: Sex:
Study design	Audit ☐ Cross-sectional ☐ Prospective ☐ Retrospective ☐ RCT ☐ Other ☐ _____
Antihypertensive drug class	
Biochemical tests monitored	
Results	*Proportion of patients with:* Baseline testing: Follow-up monitoring: *Patient factors associated with monitoring:* *Frequency of monitoring:* *Additional results presented:*
Comments (e.g. references to other studies)	

Appendix 5 – Studies investigating the nature and frequency of laboratory monitoring of patients treated with antihypertensive drugs

Study	Study design/description	Drug class	What to monitor?	Results
Rhodes (1992) UK	Cross-sectional review of computerized patient records from one general practice	Diuretics	Urea, electrolytes	76/330 (23%) patients had no record or urea or electrolyte during treatment with a diuretic but not subsequently 36 (11%) had urea and electrolyte levels measured prior to starting treatment with a diuretic 66 (20%) had results recorded within the past year 158 (48%) had results recorded within the past four years
Kalra (1999) UK	Postal questionnaire to 400 GPs; audit of 1 general practice	ACE inhibitors	Renal function (baseline and follow-up monitoring)	*Questionnaire results* GPs who usually monitored renal function: - Before start of treatment* = 235/277 (85%) - After start of treatment* = 93/277 (34%) - At no stage of treatment = 42/277 (15%) - GPs who would welcome guidelines for monitoring renal function = 234 (84%) * Time frame not specified *Audit results* Patients who received renal function monitoring: - Within 3 months before start of treatment = 55/122 (45%) - At any stage before start of treatment = 60/122 (49%) - Within 3 months after start of treatment = 35/122 (29%) - At any stage after start of treatment = 76/122 (62%) Renal dysfunction (increase of >10% in creatinine concentration) was observed in 15/122 (12%) patients, of whom 11 (73%) continued treatment without further monitoring or investigation

Study	Study design/description	Drug class	What to monitor?	Results		
Hoch (2003) Israel	Retrospective review of computerized medical records pre- and post-initiation of an electronic alert for potassium testing	Diuretics	Potassium within 1 year	*Pre-intervention* 78.5% of 34284 patients had a record for at least one serum potassium concentration test within one year of treatment *Post-intervention* 81.5% of 35313 patients had a record for at least one serum potassium concentration test within one year of treatment after the initiation of a computer alert		
Hurley (2005) USA	Retrospective analysis of databases from 2 HMOs over 3 years			% of patients with at least 1 follow-up test within the year		
				1999	*2000*	*2001*
		ACE inhibitors	Creatinine	62%	66%	68%
			Potassium	58%	61%	62%
		Diuretics	Creatinine	60%	65%	67%
			Potassium	65%	67%	67%

Study	Study design/description	Drug class	What to monitor?	Results
Simon (2005) USA	Cross-sectional analysis of databases from 10 HMOs over a 30-month-period in patients aged 65 and older	ACE inhibitors, AT-II receptor antagonists, Diuretics	Baseline creatinine or potassium testing (defined as a test 180 days before prescription and up to 14 days after)	*Patients treated with ACE inhibitors* - 6798/20 445 (33.3%, 95% CI 32.6–33.9) had no baseline creatinine or potassium testing - There was no sex difference in baseline testing (OR 1.05, 95% CI 0.99–1.11) *Patients treated with AT-II receptor antagonists* - 1 080/3 858 (28.0%, 95% CI 26.6–29.4) had no baseline creatinine or potassium testing - There was no sex difference in baseline testing (OR 0.98, 95% CI 0.84–1.14) *Patients treated with diuretics* - 11 777/35 707 (33.0%, 95% CI 32.5–33.5) had no baseline creatinine or potassium testing - Women were more likely to have no baseline testing (OR 1.25, 95% CI 1.20–1.31) - Those with no or one co-morbid condition were more likely to have no baseline testing (OR 3.63, 95% CI 3.33–3.96)
Clayton (2006b) UK	Cross-sectional study of the electronic prescribing and laboratory records of 6 UK general practices	Thiazide diuretics	Sodium and potassium electronic records	- 488/2942 (16.6%) had a sodium and/or potassium record within 2 year prior to thiazide diuretic initiation - 951/2942 (32.3%) had a sodium and/or potassium record in the two year whilst prescribed a thiazide diuretic - 140 (4.7%) patients had electrolytes checked both prior to and during treatment - Patient covariates associated with monitoring during treatment were male sex, age, and increasing number of prescriptions

Study	Study design/description	Drug class	What to monitor?	Results
Sauer (2006) USA	Retrospective analysis of a large veterans' database, identifying patients who were treated with an ACE inhibitor or AT-II receptor antagonist	ACE inhibitor or AT-II receptor antagonist	Potassium and creatinine	- 569/2936 patients did not have a baseline serum potassium or creatinine test - The frequency of patients who were not monitored within the recommended interval were: - 12 weeks (50%), 4 weeks (67%), and 2 weeks (73%) - Factors associated with lack of monitoring included: fewer than 6 outpatient encounters (OR 1.9, 95% CI 1.5–2.3), and driving distance >30 miles (OR 1.19, 95% CI 1.03–1.4)

Study	Study design/description	Drug class	What to monitor?	Results
Lafata (2007) USA	Cluster-randomized study of academic detailing in primary care practices	ACE inhibitors/ AT-II receptor antagonists; Diuretics	Baseline testing of potassium (for diuretics) or potassium and creatinine during the 180 days before or 14 days after first dispensing or within 12 months after the patient's clinic's initial academic detailing visit for continuing users	Rates of monitoring pre-intervention in control group only *ACE inhibitors/ AT-II receptor antagonists* For initial medication users: 59.2% For continuing medication users: 76.4% *Diuretics* For initial medication users: 61.2% For continuing medication users: 77.7% Rates of monitoring post-intervention in control group only* *ACE inhibitors/ AT-II receptor antagonists* For initial medication users: 51.9% For continuing medication users: 86.9% *Diuretics* For initial medication users: 62.8% For continuing medication users: 86.9% * adjusted for patients' clinic preintervention laboratory monitoring, age, gender, insurance sponsorship, Charlson Comorbidity Score, number of dispensings, number of visits to primary care provided, number of visits to other primary care physicians, percent below federal poverty line in residential census block, percent with less than high school education in residential census block, and organization

Study	Study design/description	Drug class	What to monitor?	Results
McAlister (2007) Canada	Retrospective analysis of an administrative health database over an eight-year period	Antihypertensive monotherapy (ACE inhibitor, AT-II receptor antagonist, beta-blocker, Ca-channel blocker, thiazide diuretic)	Renal function, electrolytes	- 67 879/164 413 (41%) had at least one laboratory test in 6 months before initiation of treatment - 79 985/164 413 (49%) had at least one laboratory monitoring test during 1 year follow-up In patients treated with thiazide diuretics - 21% had electrolytes checked in 6 months before first prescription; 38% in 1 year after starting treatment - 32% had renal function checked in 6 months before first prescription; 41% in 1 year after starting treatment In patients treated with newer agents (ACE inhibitors, AT-II receptor antagonists, Ca-channel blockers) - 23% had electrolytes checked in 6 months before first prescription; 31% in 1 year after starting treatment - 36% had renal function checked in 6 months before first prescription; 42% in 1year after starting treatment Compared with patients treated with newer agents, thiazide diuretic-treated patients were more likely to have their serum electrolytes monitored (OR 1.38, 95% CI 1.35–1.41) and less likely to have renal function monitored (OR 0.95, 95% CI 0.93–0.97) at least once during follow-up

Study	Study design/description	Drug class	What to monitor?	Results
Raebel (2007a) USA	Retrospective analysis of data from 10 HMO databases	Spironolactone	Creatinine and potassium within 1 year	- Both serum creatinine and potassium were evaluated at least once in 1632/2257 (72.3%) patients - 91/2257 (4.0%) patients underwent serum creatinine testing but not potassium testing - 111/2257 (4.9%) patients underwent serum potassium testing but no creatinine testing Characteristics associated with laboratory monitoring within 1 year: Male sex (adjusted OR 1.25, 95% CI 1.01–1.54); age in increments of 10 years (OR 1.28, 95% CI 1.17–1.41); outpatient visits in increments of 5 (OR 1.31, 95% CI 1.19–1.44); ACE inhibitors/AT-II receptor antagonists (OR 2.23, 95% CI 1.74–2.87); digoxin (OR 2.10, 95% CI 1.48–2.98); other diuretics (OR 1.96, 95% CI 1.51–2.54); diabetes mellitus (OR 1.63, 95% CI 1.31–2.03)

Study	Study design/description	Drug class	What to monitor?	Results
Raebel (2007b) USA	Retrospective analysis of data from 10 HMO databases	ACE inhibitors, AT-II receptor antagonists	Creatinine and potassium within 1 year	In patients treated with an ACE inhibitor: - 31 909/47 291 (67.5%) had both tests evaluated within 1 year In patients treated with an AT-II receptor antagonist - 3063/1494 (74.3%) had both tests evaluated within 1 year In all patients - Both serum creatinine and potassium were evaluated at least once in 36 185/52 096 (68.4%) patients - 3117/52 906 (5.9%) patients underwent serum creatinine testing but not potassium testing - 1001/52 906 (1.9%) patients underwent serum potassium testing but no creatinine testing Characteristics associated with laboratory monitoring within 1 year for all patients: Male sex (OR 1.06; 95% CI 1.02–1.11); age (80+ compared with <50) (OR 2.10, 95% CI 1.93–2.28); outpatient visits (>9 compared to ≤9) (OR 1.46, 95% CI 1.39–1.54); any hospitalization (OR 1.15, 95% CI 1.06–1.25); digoxin (OR 1.15, 95% CI 1.01–1.30); potassium supplements (OR 2.01, 95% CI 1.84–2.20); diuretics (OR 1.54, 95% CI 1.47–1.61); diabetes mellitus (OR 1.68, 95% CI 1.61–1.75); heart failure (OR 1.73, 95% CI 1.57–1.90); chronic kidney disease (OR 2.95, 95% CI 2.48–3.51)
Besançon (2008) France	Retrospective review of anonymous computerized healthcare records from a health insurance system	Spironolactone and ACE inhibitor	Serum potassium and creatinine	1083 (30%) of patients had at least one measurement of potassium or creatinine in the six months prior to treatment 15% underwent either serum potassium or serum creatinine measurements but not both 34% did not undergo any laboratory tests Only 51% underwent the minimal biological monitoring defined in the reference system.

Study	Study design/description	Drug class	What to monitor?	Results
Géradin-Marais (2008) France	One-year retrospective analysis of a national health insurance database in patients aged 75 and older	Diuretics (loop diuretics, thiazide diuretics, potassium-sparing agents, combination regimen)	Serum chemistry monitoring (defined as urea and creatinine clearance; urea, creatinine, potassium, sodium, chloride or carbon dioxide serum levels, plasma proteins)	- A total of 11 315 patients, aged 75 years or more, were dispensed at least 12 monthly diuretic prescriptions - 8513/11 315 (75%) had at least one monitoring test within one year of treatment Patient characteristics associated with monitoring: Patient age ≥85 years (OR 1.63, 1.36–1.97); female sex (OR 1.28, 95% CI 1.15–1.42); thiazide diuretics (compared with loop diuretics) (OR 1.54, 95% CI 1.21–1.95); serious disease (OR 1.70, 95% CI 1.55–1.87)
Raebel (2010) USA	Retrospective analysis of data from 3 HMOs in diabetic patients aged 18 and older newly treated with antihypertensive therapy	ACE inhibitors, angiotensin-II receptor blockers, spironolactone	Serum potassium	19 391/27 355 (71%) of patients had at least one serum potassium test during treatment. 244/27 335 (0.9%) had a baseline potassium test within 30 days prior to initiating treatment. Patients who had at least one potassium test were less likely to experience serious hyperkalaemia (≥6.0mmol/l), or hyperkalaemia-related adverse events such as hospital admission, emergency department visit, or death. Unadjusted RR 1.29, 95% CI 0.87–1.90 Adjusted RR* 0.50, 95% CI 0.37–0.66 (*adjusted for potential confounders of monitoring) In the subgroup of patients with chronic disease, monitoring further reduced the risk of harm. Unadjusted RR 0.28, 95% CI 0.14–0.57 Adjusted RR* 0.19, 95% CI 0.11–0.36 (*adjusted for potential confounders of monitoring)

Appendix 6 – Studies investigating the nature and frequency of laboratory monitoring excluded from the systematic review

Study	Study design/description	Drug class	What to monitor?	Results
Excluded because discrete data on only the antihypertensive medication could not be obtained				
Feldstein (2006) USA	Cluster-randomized trial comparing 3 interventions to usual care in 15 primary care clinics in one HMO	ACE inhibitors, AT-II receptor antagonists, diuretics (as well as 8 other drug classes)	Baseline and follow-up creatinine, potassium	*Proportion of patients with baseline laboratory testing* Usual care 60.4% Intervention 1 (electronic medical record) 59.7% Intervention 2 (automated voice messaging) 53.2% Intervention 3 (pharmacy team) 59.1% *Proportion of patients with follow-up laboratory testing within 25 days* Usual care 22.4% Intervention 1 (electronic medical record) 48.5% Intervention 2 (automated voice messaging) 66.3% Intervention 3 (pharmacy team) 82.0%
Steele (2005) USA	Pre- and post-intervention analysis	ACE inhibitors, thiazide diuretics (and additional non-antihypertensive drug classes)	Follow-up creatinine and potassium	Monitoring following an intervention using an electronic alert Intervention / Pre-intervention / Post-intervention / P value No message displayed 17% 16.2% 0.38 Message displayed 38.5% 51.1% <0.001 Message displayed: "Abnormal Labs" 33.8% 41.7% 0.077 Message displayed: "No Labs Available" 43.0% 62.0% <0.001

Study	Study design/description	Drug class	What to monitor?	Results
Excluded because the denominator presented was the number of prescriptions and not the number of patients				
Raebel (2005) USA	Retrospective analysis of patients from 10 HMOs who were newly prescribed treatment	ACE inhibitors, AT-II receptor antagonists, diuretics	Baseline testing of creatinine or potassium (180 days before and 14 days after index prescription)	*ACE inhibitors* - 39.2% of 48 682 dispensings had no creatinine or potassium test - 32.3% of dispensing had no creatinine test *AT-II receptor antagonists* - 34.1% of 8731 dispensings had no creatinine or potassium test - 29.6% of dispensings had no creatinine test *Diuretics* - 39.5% of 78 903 dispensings had no creatinine or potassium test - 34.4% of dispensing did had no creatinine test

Study	Study design/description	Drug class	What to monitor?	Results
Palen (2006) USA	Randomized intervention study of a computerized physician order entry system in a US managed care organization over 1 year	ACE inhibitors, AT-II receptor antagonists, diuretics	Creatinine and potassium	*Control group* - 47.5% prescription orders had a creatinine and/or potassium record in the 180 days prior to and 14 days after the initial ACE inhibitor treatment - 52.7% prescription orders had a creatinine and/or potassium record in the 180 days prior to and 14 days after the initial AT-II receptor antagonist treatment - 45.6% prescription orders had a creatinine and/or potassium record in the 180 days prior to and 14 days after the diuretic treatment *Intervention group (non-intrusive laboratory monitoring electronic alert)* - 47.0% prescription orders had a creatinine and/or potassium record in the 180 days prior to and 14 days after the initial ACE inhibitor treatment - 52.0% prescription orders had a creatinine and/or potassium record in the 180 days prior to and 14 days after the initial AT-II receptor antagonist treatment - 44.0%% prescription orders had a creatinine and/or potassium record in the 180 days prior to and 14 days after the diuretic treatment No significant differences were observed in the rates of monitoring in group where electronic alerts reminded the physician of the need for laboratory monitoring.

Study	Study design/description	Drug class	What to monitor?	Results
Matheny (2008) USA	Randomized study of electronic reminders in a cohort of patients registered in a database over 6 months	AT-II receptor antagonists, potassium-sparing thiazide diuretics, diuretics, ACE inhibitors	Creatinine, potassium	Control group - 3.2% of GP visits of patients treated with an AT-II receptor antagonist had no record of a creatinine test in the 365 days prior - 3.2% of GP visits of patients treated with potassium-sparing diuretic had no record of a potassium test in the 365 days prior - 3.5% of GP visits of patients treated with a thiazide diuretic had no record of a potassium test in the 365 days prior - 2.9% of GP visits of patients treated with an ACE inhibitor had no record of a potassium test in the 365 days prior Intervention group (electronic reminders of laboratory monitoring) - 4.1% of GP visits of patients treated with an AT-II receptor antagonist had no record of a creatinine test in the 365 days prior - 2.5% of GP visits of patients treated with potassium-sparing diuretic had no record of a potassium test in the 365 days prior - 3.1% of GP visits of patients treated with a thiazide diuretic had no record of a potassium test in the 365 days prior - 5.2% of GP visits of patients treated with an ACE inhibitor had no record of a potassium test in the 365 days prior No statistically significant improvement in any of the monitoring was observed with the intervention

Appendix 7 – Read/OXMIS codes for hypertension

Code	Read/OXMIS name	Code	Read/OXMIS name
403 AH	RENOVASCULAR HYPERTENSION	9OI..00	Hypertension monitoring admin.
662B.00	O/E - initial high BP	Y0601JA	HYPERTENSION CLINIC ATTENDANCE
662G.00	Hypertensive treatm.changed	4000	MALIGNANT HYPERTENSION
9OI..11	Hypertension clinic admin.	401 AR	HYPERTENSION ARTERIOSCLEROTIC
G2...11	BP - hypertensive disease	662F.00	Hypertension treatm. started
G201.00	Benign essential hypertension	9OIA.11	Hypertension monitored
G232.00	Hypertensive heart&renal dis wth (congestive) heart failure	G20z.00	Essential hypertension NOS
G241z00	Secondary benign hypertension NOS	G210100	Malignant hypertensive heart disease with CCF
Y060 JA	HYPERTENSION CLINIC	G211z00	Benign hypertensive heart disease NOS
Y100 NH	NURSE HYPERTENSION ASSESSMENT	G222.00	Hypertensive renal disease with renal failure
401 P	HYPERTENSION ARTERIAL PRIMARY	G240z00	Secondary malignant hypertension NOS
9N1y200	Seen in hypertension clinic	G24zz00	Secondary hypertension NOS
G210z00	Malignant hypertensive heart disease NOS	G672.11	Hypertensive crisis
Gyu2000	[X]Other secondary hypertension	403 AA	RENAL HYPERTENSIVE DISEASE
L0010EM	REFERRED TO HYPERTENSION CLINIC	9OI2.00	Refuses hypertension monitor.
4003AA	HYPERTENSIVE RENAL DISEASE	9OIA.00	Hypertension monitor.chck done
401 R	HYPERTENSION CRISIS	G202.00	Systolic hypertension
6627.00	Good hypertension control	G20z.11	Hypertension NOS
9OI1.00	Attends hypertension monitor.	G21zz00	Hypertensive heart disease NOS
G210000	Malignant hypertensive heart disease without CCF	G23..00	Hypertensive heart and renal disease
G211100	Benign hypertensive heart disease with CCF	G2z..00	Hypertensive disease NOS
G21z100	Hypertensive heart disease NOS with CCF	Gyu2100	[X]Hypertension secondary to other renal disorders
G221.00	Benign hypertensive renal disease	401 A	HYPERTENSION
G24z.00	Secondary hypertension NOS	401 AB	HYPERTENSION BORDERLINE
G2y..00	Other specified hypertensive disease	401 AW	HIGH BLOOD PRESSURE
Gyu2.00	[X]Hypertensive diseases	401 BN	HYPERTENSION BENIGN
L0010AM	SEEN IN HYPERTENSION CLINIC	401 C	HYPERTENSION ESSENTIAL
403 NC	HYPERTENSION IMPAIRED RENAL FUNCTION	401 DC	HYPERTENSION DIASTOLIC
2466.00	O/E - BP reading raised	403	RENAL HYPERTENSION
662O.00	On treatment for hypertension	401 AC	HYPERTENSION ACCELERATED
662P.00	Hypertension monitoring	401 AK	OBSERVATION ONLY HYPERTENSION

Code	Read/OXMIS name	Code	Read/OXMIS name
8B26.00	Antihypertensive therapy	G240.00	Secondary malignant hypertension
G2...00	Hypertensive disease	2465.00	O/E - BP borderline raised
G20..11	High blood pressure	G200.00	Malignant essential hypertension
G210.00	Malignant hypertensive heart disease	G21..00	Hypertensive heart disease
G21z.00	Hypertensive heart disease NOS	G211.00	Benign hypertensive heart disease
G231.00	Benign hypertensive heart and renal disease	G22..00	Hypertensive renal disease
G240000	Secondary malignant renovascular hypertension	G230.00	Malignant hypertensive heart and renal disease
G241000	Secondary benign renovascular hypertension	G244.00	Hypertension secondary to endocrine disorders
4003	HYPERTENSIVE NEPHROPATHY	401	BENIGN HYPERTENSION
G220.00	Malignant hypertensive renal disease	401 EL	HYPERTENSION ARTERIAL SYSTEMIC
G22z.00	Hypertensive renal disease NOS	402 C	HYPERTENSION CONGESTIVE HEART FAILURE
G233.00	Hypertensive heart and renal disease with renal failure	403 NG	HYPERTENSIVE GLOMERULOSCLEROSIS
G23z.00	Hypertensive heart and renal disease NOS	8HT5.00	Referral to hypertension clinic
G241.00	Secondary benign hypertension	9N03.00	Seen in hypertension clinic
G24z000	Secondary renovascular hypertension NOS	401 BM	HYPERTENSION ON TREATMENT
G672.00	Hypertensive encephalopathy	401 S	HYPERTENSION SYSTOLIC
3055CE	HYPERTENSIVE SPASM CARDIAC SPHINCTER	402	HYPERTENSIVE HEART DISEASE
401 AT	HYPERTENSION ATHEROSCLEROTIC	9N4L.00	DNA - Did not attend hypertension clinic
401 PA	HYPERTENSION PRIMARY	662d.00	Hypertension annual review
2467.00	O/E - BP reading very high	662c.00	Hypertension six month review
6628.00	Poor hypertension control	662b.00	Moderate hypertension control
G20..00	Essential hypertension	8BL0.00	Patient on maximal tolerated antihypertensive therapy
G21z000	Hypertensive heart disease NOS without CCF	66b2.00	Hypertension monitoring not required
G22z.11	Renal hypertension	1JD..00	Suspected hypertension
G24..00	Secondary hypertension	8CR4.00	Hypertension clinical management plan
401 E	HYPERTENSION ARTERIAL	G8y3.00	Chronic peripheral venous hypertension
403 NF	HYPERTENSION RENAL INSUFFICIENCY	G21z011	Cardiomegaly - hypertensive
662..12	Hypertension monitoring	8I3N.00	Hypertension treatment refused
662C.00	O/E - check high BP	G211000	Benign hypertensive heart disease without CCF

Appendix 8 – Read/OXMIS codes for baseline smoking

Code	Read/OXMIS name	Code	Read/OXMIS name
1373.00	Light smoker - 1-9 cigs/day	9OO2.00	Refuses stop smoking monitor
137Q.11	Smoking restarted	T5113	SMOKER (15 PER DAY)
137R.00	Current smoker	T512	SMOKER PIPE
9OO1.00	Attends stop smoking monitor.	1377.00	Ex-trivial smoker (
T5091HS	EX HEAVY SMOKER	1379.00	Ex-moderate smoker (10-19/day)
T5115M	SMOKER MILD (5 OR LESS PER DAY)	137A.00	Ex-heavy smoker (20-39/day)
T5117	SMOKER (30 PER DAY)	9OO5.00	Stop smoking monitor 2nd lettr
T513	SMOKER CIGARS	9OO6.00	Stop smoking monitor 3rd lettr
L5091S	SMOKING STARTED	9OO7.00	Stop smoking monitor verb.inv.
T5093N	SMOKED NEVER	9OO9.00	Stop smoking monitoring delete
Y060 J1	STOP SMOKING GROUP	T509	SMOKER
Y060 JJ	CLINIC ANTI-SMOKING	T510 SE	SMOKING EXCESSIVE
137B.00	Ex-very heavy smoker (40+/day)	T5114	SMOKER (10 PER DAY)
H310100	Smokers' cough	1371.00	Never smoked tobacco
T5091ES	FORMER SMOKER	1372.00	Trivial smoker - < 1 cig/day
T5092	SMOKING ADVISED TO STOP	1376.00	Very heavy smoker - 40+cigs/d
T5093	SMOKER NON	137H.00	Pipe smoker
T510	EXCESSIVE SMOKING	137K.00	Stopped smoking
T5112	SMOKER (20 PER DAY)	9OO..11	Stop smoking clinic admin.
5020M	SMOKERS' THROAT	9OO..12	Stop smoking monitoring admin.
137..11	Smoker - amount smoked	T509 SR	SMOKING RESTARTED
1371.11	Non-smoker	137C.00	Keeps trying to stop smoking
137F.00	Ex-smoker - amount unknown	137P.11	Smoker
137L.00	Current non-smoker	9OOA.00	Stop smoking monitor.chck done
9OO..00	Anti-smoking monitoring admin.	T5090XC	SMOKER CIGARETTES
T5092SA	SMOKING WANTS TO STOP	T510 HS	HEAVY SMOKER (20-PLUS PER DAY)
137M.00	Rolls own cigarettes	T5115	SMOKER (LESS THAN 10 PER DAY)
137N.00	Ex pipe smoker	8CAL.00	Smoking cessation advice
13WF400	Passive smoking risk	8HTK.00	Referral to stop-smoking clinic
9OO8.00	Stop smoking monitor phone inv	137V.00	Smoking reduced
T5090OR	SMOKER OWN ROLLED	13p..00	Smoking cessation milestones
T5092S	SMOKING WISHES TO STOP	137X.00	Cigarette consumption
T511	SMOKER MODERATE (LESS THAN 20 PER DAY)	13p1.00	Smoking status at 4 weeks
Y0601KA	SMOKING CLINIC ATTENDANCE	137Y.00	Cigar consumption
5287MK	SMOKERS' MOUTH PATCHES	13p0.00	Negotiated date for cessation of smoking
1372.11	Occasional smoker	13p2.00	Smoking status between 4 and 52 weeks
1375.00	Heavy smoker - 20-39 cigs/day	137b.00	Ready to stop smoking
1378.00	Ex-light smoker (1-9/day)	137c.00	Thinking about stopping smoking
137J.00	Cigar smoker	13p4.00	Smoking free weeks
137Q.00	Smoking started	9N4M.00	DNA - Did not attend smoking cessation clinic
137T.00	Date ceased smoking	8H7i.00	Referral to smoking cessation advisor
9OO4.00	Stop smoking monitor 1st lettr	13p3.00	Smoking status at 52 weeks
9OOZ.00	Stop smoking monitor admin.NOS	13p5.00	Smoking cessation programme start date
T5091	STOPPED SMOKING	137g.00	Cigarette pack-years
137P.00	Cigarette smoker	137d.00	Not interested in stopping smoking

Code	Read/OXMIS name	Code	Read/OXMIS name
T510 SH	SMOKER HEAVY (20-PLUS PER DAY)	9N2k.00	Seen by smoking cessation advisor
T5116	SMOKER(OCCASIONAL)	137e.00	Smoking restarted
1374.00	Moderate smoker - 10-19 cigs/d	137f.00	Reason for restarting smoking
137G.00	Trying to give up smoking	67H1.00	Lifestyle advice regarding smoking
137O.00	Ex cigar smoker	ZRh4.00	Reasons for smoking scale
137S.00	Ex smoker	13WK.00	No smokers in the household
8I6H.00	Smoking review not indicated	745H.00	Smoking cessation therapy
ZRh4.11	RFS - Reasons for smoking scale	ZRaM.00	Motives for smoking scale

Appendix 9 – Read/OXMIS codes for baseline diabetes mellitus

Code	Read/OXMIS name	Code	Read/OXMIS name
R054200	[D]Gangrene of toe in diabetic	L2500GC	DIABETES - GOOD CONTROL
R054300	[D]Widespread diabetic foot gangrene	C10yz00	Diabetes mellitus NOS with other specified manifestation
ZV65312	[V]Dietary counselling in diabetes mellitus	C107z00	Diabetes mellitus NOS with peripheral circulatory disorder
Cyu2.00	[X]Diabetes mellitus	C10zz00	Diabetes mellitus NOS with unspecified complication
Kyu0300	[X]Glomerular disorders in diabetes mellitus	C107.11	Diabetes mellitus with gangrene
Cyu2100	[X]Malnutrit-relat diabetes mellitus with other spec comps	C102.00	Diabetes mellitus with hyperosmolar coma
Cyu2200	[X]Malnutrit-related diabetes mellitus with unspec complics	C101.00	Diabetes mellitus with ketoacidosis
Cyu2000	[X]Other specified diabetes mellitus	C103.00	Diabetes mellitus with ketoacidotic coma
Cyu2300	[X]Unspecified diabetes mellitus with renal complications	C106.00	Diabetes mellitus with neurological manifestation
250 AB	ABSCESS DIABETIC	C106.12	Diabetes mellitus with neuropathy
F372000	Acute painful diabetic neuropathy	C100.00	Diabetes mellitus with no mention of complication
8H2J.00	Admit diabetic emergency	C105.00	Diabetes mellitus with ophthalmic manifestation
F420300	Advanced diabetic maculopathy	C10y.00	Diabetes mellitus with other specified manifestation
F420500	Advanced diabetic retinal disease	C107.00	Diabetes mellitus with peripheral circulatory disorder
TJ23z00	Adverse reaction to insulins and antidiabetic agents NOS	C106.13	Diabetes mellitus with polyneuropathy
66AT.00	Annual diabetic blood test	C104.00	Diabetes mellitus with renal manifestation
F372200	Asymptomatic diabetic neuropathy	C10z.00	Diabetes mellitus with unspecified complication
9NM0.00	Attending diabetes clinic	C106100	Diabetes mellitus, adult onset, + neurological manifestation
F171100	Autonomic neuropathy due to diabetes	C105100	Diabetes mellitus, adult onset, + ophthalmic manifestation
F420000	Background diabetic retinopathy	C10z100	Diabetes mellitus, adult onset, + unspecified complication
M037200	Cellulitis in diabetic foot	C100100	Diabetes mellitus, adult onset, no mention of complication
250 M	CHARCOT'S DIABETIC ARTHROPATHY	C102100	Diabetes mellitus, adult onset, with hyperosmolar coma
F372100	Chronic painful diabetic neuropathy	C101100	Diabetes mellitus, adult onset, with ketoacidosis
250 H	COMA DIABETIC	C103100	Diabetes mellitus, adult onset, with ketoacidotic coma
66AN.00	Date diabetic treatment start	C104100	Diabetes mellitus, adult onset, with renal manifestation
250 AN	DIABETES	C107200	Diabetes mellitus, adult with gangrene
C10y100	Diabetes mellitus, adult, + other specified manifestation	66A..00	Diabetic monitoring
250 JE	DIABETIC ACETONAEMIA	66Al.00	Diabetic monitoring - higher risk albumin excretion
250 JA	DIABETIC ACIDOSIS	66Ak.00	Diabetic monitoring - lower risk albumin excretion

Code	Read/OXMIS name	Code	Read/OXMIS name
250 AT	DIABETIC AMYOTROPHY	66AZ.00	Diabetic monitoring NOS
F381311	Diabetic amyotrophy	F345000	Diabetic mononeuritis multiplex
C106.11	Diabetic amyotrophy	F35z000	Diabetic mononeuritis NOS
66AS.00	Diabetic annual review	F3y0.00	Diabetic mononeuropathy
F464000	Diabetic cataract	C104.11	Diabetic nephropathy
250 CT	DIABETIC CATARACT	250 N	DIABETIC NEPHROPATHY
N030100	Diabetic Charcot arthropathy	F372.12	Diabetic neuropathy
N030000	Diabetic cheiroarthropathy	66A3.00	Diabetic on diet only
N030011	Diabetic cheiropathy	66A5.00	Diabetic on insulin
2500W	DIABETIC CLINIC	66AV.00	Diabetic on insulin and oral treatment
8A12.00	Diabetic crisis monitoring	66A4.00	Diabetic on oral treatment
250 DR	DIABETIC DIARRHOEA	9OLD.00	Diabetic patient unsuitable for digital retinal photography
13B1.00	Diabetic diet	G73y000	Diabetic peripheral angiopathy
66AY.00	Diabetic diet - good compliance	66Ac.00	Diabetic peripheral neuropathy screening
66Aa.00	Diabetic diet - poor compliance	F372.11	Diabetic polyneuropathy
68AB.00	Diabetic digital retinopathy screening offered	6761.00	Diabetic pre-pregnancy counselling
66AG.00	Diabetic drug side effects	F420.00	Diabetic retinopathy
66Ab.00	Diabetic foot examination	8HBG.00	Diabetic retinopathy 12 month review
8I3W.00	Diabetic foot examination declined	8HBH.00	Diabetic retinopathy 6 month review
8I6G.00	Diabetic foot examination not indicated	F420z00	Diabetic retinopathy NOS
66AW.00	Diabetic foot risk assessment	68A7.00	Diabetic retinopathy screening
250 LG	DIABETIC GLOMERULOSCLEROSIS	8I6F.00	Diabetic retinopathy screening not indicated
13AB.00	Diabetic lipid lowering diet	8I3X.00	Diabetic retinopathy screening refused
F420400	Diabetic maculopathy	8A13.00	Diabetic stabilisation
66AH.00	Diabetic treatment changed	250 HC	HYPOGLYCAEMIC COMA DIABETIC
13AC.00	Diabetic weight reducing diet	250 ED	HYPOGLYCAEMICS ORAL DIABETES
66AL.00	Diabetic-uncooperative patient	9999DM	IATROGENIC DIABETES MELLITUS
8HLE.00	Diabetology D.V. done	C108.11	IDDM-Insulin dependent diabetes mellitus
8HKE.00	Diabetology D.V. requested	66AA.11	Injection sites - diabetic
9N4p.00	Did not attend diabetic retinopathy clinic	C108900	Insulin dependent diabetes maturity onset
ZC2C800	Dietary advice for diabetes mellitus	C100011	Insulin dependent diabetes mellitus
ZC2C900	Dietary advice for type I diabetes	C10E.12	Insulin dependent diabetes mellitus
ZC2CA00	Dietary advice for type II diabetes	C108.00	Insulin dependent diabetes mellitus
250 DC	DIETARY CONTROL DIABETES	C10E812	Insulin dependent diabetes mellitus - poor control
9N4I.00	DNA - Did not attend diabetic clinic	C108800	Insulin dependent diabetes mellitus - poor control
2G5C.00	Foot abnormality - diabetes related	C108H00	Insulin dependent diabetes mellitus with arthropathy
2G51000	Foot abnormality - diabetes related	C108F00	Insulin dependent diabetes mellitus with diabetic cataract

Code	Read/OXMIS name	Code	Read/OXMIS name
250 GA	GANGRENE DIABETIC	C108600	Insulin dependent diabetes mellitus with gangrene
44V3.00	Glucose tol. test diabetic	C108E00	Insulin dependent diabetes mellitus with hypoglycaemic coma
14F4.00	H/O: Admission in last year for diabetes foot problem	C108B00	Insulin dependent diabetes mellitus with mononeuropathy
1434.00	H/O: diabetes mellitus	C10E312	Insulin dependent diabetes mellitus with multiple complicat
66A8.00	Has seen dietician - diabetes	C108300	Insulin dependent diabetes mellitus with multiple complicatn
42W..00	Hb. A1C - diabetic control	C108D00	Insulin dependent diabetes mellitus with nephropathy
42WZ.00	Hb. A1C - diabetic control NOS	C108C00	Insulin dependent diabetes mellitus with polyneuropathy
42c..00	HbA1 - diabetic control	C108700	Insulin dependent diabetes mellitus with retinopathy
F420800	High risk non proliferative diabetic retinopathy	C108500	Insulin dependent diabetes mellitus with ulcer
F420700	High risk proliferative diabetic retinopathy	C109J11	Insulin treated non-insulin dependent diabetes mellitus
250 NH	HYPEROSMOLAR DIABETIC STATE	C109J00	Insulin treated Type 2 diabetes mellitus
C10FK00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus	C10FJ00	Insulin treated Type 2 diabetes mellitus
C109K00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus	C109J12	Insulin treated Type II diabetes mellitus
250 E	HYPOGLYCAEMIA IN DIABETES MELLITUS	C10FJ11	Insulin treated Type II diabetes mellitus
C108200	Insulin-dependent diabetes mellitus with neurological comps	M271200	Mixed diabetic ulcer - foot
C108100	Insulin-dependent diabetes mellitus with ophthalmic comps	F381300	Myasthenic syndrome due to diabetic amyotrophy
C108000	Insulin-dependent diabetes mellitus with renal complications	K01x100	Nephrotic syndrome in diabetes mellitus
C108A00	Insulin-dependent diabetes without complication	M271100	Neuropathic diabetic ulcer - foot
M271000	Ischaemic ulcer diabetic foot	250 F	NEUROPATHY DIABETIC
250 JK	KETOACIDOSIS DIABETIC	C109.11	NIDDM - Non-insulin dependent diabetes mellitus
250 JL	KETOSIS DIABETIC	250 AA	NIDDM (NON-INSULIN DEPENDENT DIABETES)
2500AH	LATENT DIABETES	F420600	Non proliferative diabetic retinopathy
C10M.00	Lipoatrophic diabetes mellitus	C109E00	Non-insulin depend diabetes mellitus with diabetic cataract
8HME.00	Listed for Diabetology admissn	C109700	Non-insulin dependant diabetes mellitus - poor control
C10A.00	Malnutrition-related diabetes mellitus	C100112	Non-insulin dependent diabetes mellitus
C10A000	Malnutrition-related diabetes mellitus with coma	C109G00	Non-insulin dependent diabetes mellitus with arthropathy
C10A100	Malnutrition-related diabetes mellitus with ketoacidosis	C109500	Non-insulin dependent diabetes mellitus with gangrene
C10A600	Malnutrition-related diabetes mellitus with multiple comps	C109D00	Non-insulin dependent diabetes mellitus with hypoglyca coma
C10A200	Malnutrition-related diabetes mellitus with renal complicatn	C109A00	Non-insulin dependent diabetes mellitus with mononeuropathy
C10A700	Malnutrition-related diabetes mellitus without complications	C109C00	Non-insulin dependent diabetes mellitus with nephropathy

Code	Read/OXMIS name	Code	Read/OXMIS name
C10A400	Malnutrition-related diabetes mellitus wth neuro complicatns	C109B00	Non-insulin dependent diabetes mellitus with polyneuropathy
C10A500	Malnutritn-relat diabetes melitus wth periph circul complctn	C109400	Non-insulin dependent diabetes mellitus with ulcer
C10AX00	Malnutrit-relat diabetes mellitus with other spec comps	C109.00	Non-insulin-dependent diabetes mellitus
C10AW00	Malnutrit-related diabetes mellitus with unspec complics	C109300	Non-insulin-dependent diabetes mellitus with multiple comps
C10A300	Malnutrit-related diabetes mellitus wth ophthalmic complicat	C109200	Non-insulin-dependent diabetes mellitus with neuro comps
C100111	Maturity onset diabetes	C109100	Non-insulin-dependent diabetes mellitus with ophthalm comps
250 AM	MATURITY ONSET DIABETES (MELLITUS)	C109000	Non-insulin-dependent diabetes mellitus with renal comps
C10C.11	Maturity onset diabetes in youth	C109600	Non-insulin-dependent diabetes mellitus with retinopathy
C10D.11	Maturity onset diabetes in youth type 2	C109900	Non-insulin-dependent diabetes mellitus without complication
250 AK	MATURITY ONSET DIABETES MELLITUS INSULIN	8H3O.00	Non-urgent diabetic admission
250 AL	MATURITY ONSET DIABETES(MELLITUS) NON-IN	2BBL.00	O/E - diabetic maculopathy present both eyes
2G5W.00	O/E - left chronic diabetic foot ulcer	C105y00	Other specified diabetes mellitus with ophthalmic complicatn
2G5L.00	O/E - Left diabetic foot - ulcerated	C10yy00	Other specified diabetes mellitus with other spec comps
2G5K.00	O/E - Left diabetic foot at high risk	C107y00	Other specified diabetes mellitus with periph circ comps
2G5I.00	O/E - Left diabetic foot at low risk	C104y00	Other specified diabetes mellitus with renal complications
2G5J.00	O/E - Left diabetic foot at moderate risk	C10zy00	Other specified diabetes mellitus with unspecified comps
2G5B.00	O/E - Left diabetic foot at risk	7276.00	Pan retinal photocoagulation for diabetes
2BBQ.00	O/E - left eye background diabetic retinopathy	93C4.00	Patient consent given for addition to diabetic register
2BBX.00	O/E - left eye diabetic maculopathy	9360.00	Patient held diabetic record issued
2BBS.00	O/E - left eye preproliferative diabetic retinopathy	679R.00	Patient offered diabetes structured education programme
2BBV.00	O/E - left eye proliferative diabetic retinopathy	8BL2.00	Patient on maximal tolerated therapy for diabetes
2BBl.00	O/E - left eye stable treated prolif diabetic retinopathy	ZRbH.00	Perceived control of insulin-dependent diabetes
2G5V.00	O/E - right chronic diabetic foot ulcer	F372.00	Polyneuropathy in diabetes
2G5H.00	O/E - Right diabetic foot - ulcerated	250 HP	PRECOMA DIABETIC
2G5G.00	O/E - Right diabetic foot at high risk	L180500	Pre-existing diabetes mellitus, insulin-dependent
2G5E.00	O/E - Right diabetic foot at low risk	L180600	Pre-existing diabetes mellitus, non-insulin-dependent
2G5F.00	O/E - Right diabetic foot at moderate risk	L180X00	Pre-existing diabetes mellitus, unspecified
2G5A.00	O/E - Right diabetic foot at risk	L180700	Pre-existing malnutrition-related diabetes mellitus
2BBP.00	O/E - right eye background diabetic retinopathy	F420200	Preproliferative diabetic retinopathy
2BBW.00	O/E - right eye diabetic maculopathy	8HVU.00	Private referral to diabetologist
2BBR.00	O/E - right eye preproliferative diabetic retinopathy	F420100	Proliferative diabetic retinopathy

Code	Read/OXMIS name	Code	Read/OXMIS name
2BBT.00	O/E - right eye proliferative diabetic retinopathy	250 PR	PRURITUS DIABETIC
2BBk.00	O/E - right eye stable treated prolif diabetic retinopathy	8H7r.00	Refer to diabetic foot screener
2BBo.00	O/E - sight threatening diabetic retinopathy	8Hl1.00	Referral for diabetic retinopathy screening
C103y00	Other specified diabetes mellitus with coma	ZL62500	Referral to diabetes nurse
C101y00	Other specified diabetes mellitus with ketoacidosis	8HTk.00	Referral to diabetic eye clinic
C108y00	Other specified diabetes mellitus with multiple comps	8HHy.00	Referral to diabetic register
C106y00	Other specified diabetes mellitus with neurological comps	8H4F.00	Referral to diabetologist
8HTi.00	Referral to multidisciplinary diabetic clinic	C10EN00	Type 1 diabetes mellitus with ketoacidotic coma
L0010EI	REFERRED TO DIABETIC CLINIC	C10EB00	Type 1 diabetes mellitus with mononeuropathy
9OL2.00	Refuses diabetes monitoring	C10E300	Type 1 diabetes mellitus with multiple complications
2BBF.00	Retinal abnormality - diabetes related	C10ED00	Type 1 diabetes mellitus with nephropathy
250 C	RETINOPATHY DIABETIC	C10E200	Type 1 diabetes mellitus with neurological complications
C10N.00	Secondary diabetes mellitus	C108212	Type 1 diabetes mellitus with neurological complications
C10G.00	Secondary pancreatic diabetes mellitus	C108J12	Type 1 diabetes mellitus with neuropathic arthropathy
C10G000	Secondary pancreatic diabetes mellitus without complication	C10EJ00	Type 1 diabetes mellitus with neuropathic arthropathy
9N2d.00	Seen by diabetologist	C10E100	Type 1 diabetes mellitus with ophthalmic complications
L0010AI	SEEN IN DIABETIC CLINIC	C10EG00	Type 1 diabetes mellitus with peripheral angiopathy
9N1v.00	Seen in diabetic eye clinic	C10EL00	Type 1 diabetes mellitus with persistent microalbuminuria
9N1i.00	Seen in diabetic foot clinic	C10EK00	Type 1 diabetes mellitus with persistent proteinuria
250 A	SUGAR DIABETES	C10EC00	Type 1 diabetes mellitus with polyneuropathy
8CP2.00	Transition of diabetes care options discussed	C10E000	Type 1 diabetes mellitus with renal complications
C108.12	Type 1 diabetes mellitus	C108012	Type 1 diabetes mellitus with renal complications
C10E.00	Type 1 diabetes mellitus	C108712	Type 1 diabetes mellitus with retinopathy
C10E800	Type 1 diabetes mellitus - poor control	C10E700	Type 1 diabetes mellitus with retinopathy
C108812	Type 1 diabetes mellitus - poor control	C10E500	Type 1 diabetes mellitus with ulcer
C10E900	Type 1 diabetes mellitus maturity onset	C10EA00	Type 1 diabetes mellitus without complication
C10EH00	Type 1 diabetes mellitus with arthropathy	C109.12	Type 2 diabetes mellitus
C10EF00	Type 1 diabetes mellitus with diabetic cataract	C10F.00	Type 2 diabetes mellitus
C10EP00	Type 1 diabetes mellitus with exudative maculopathy	C109712	Type 2 diabetes mellitus - poor control

Code	Read/OXMIS name	Code	Read/OXMIS name
C10E600	Type 1 diabetes mellitus with gangrene	C10F700	Type 2 diabetes mellitus - poor control
C10EQ00	Type 1 diabetes mellitus with gastroparesis	C109G12	Type 2 diabetes mellitus with arthropathy
C10EE00	Type 1 diabetes mellitus with hypoglycaemic coma	C10FG00	Type 2 diabetes mellitus with arthropathy
C108E12	Type 1 diabetes mellitus with hypoglycaemic coma	C10FE00	Type 2 diabetes mellitus with diabetic cataract
C10EM00	Type 1 diabetes mellitus with ketoacidosis	C109E12	Type 2 diabetes mellitus with diabetic cataract
C10FQ00	Type 2 diabetes mellitus with exudative maculopathy	C109412	Type 2 diabetes mellitus with ulcer
C10F500	Type 2 diabetes mellitus with gangrene	C10F900	Type 2 diabetes mellitus without complication
C10FR00	Type 2 diabetes mellitus with gastroparesis	C108.13	Type I diabetes mellitus
C109D12	Type 2 diabetes mellitus with hypoglycaemic coma	C10E.11	Type I diabetes mellitus
C10FD00	Type 2 diabetes mellitus with hypoglycaemic coma	C108811	Type I diabetes mellitus - poor control
C10FN00	Type 2 diabetes mellitus with ketoacidosis	C108911	Type I diabetes mellitus maturity onset
C10FP00	Type 2 diabetes mellitus with ketoacidotic coma	C108H11	Type I diabetes mellitus with arthropathy
C10FA00	Type 2 diabetes mellitus with mononeuropathy	C108F11	Type I diabetes mellitus with diabetic cataract
C10F300	Type 2 diabetes mellitus with multiple complications	C108E11	Type I diabetes mellitus with hypoglycaemic coma
C10FC00	Type 2 diabetes mellitus with nephropathy	C10EM11	Type I diabetes mellitus with ketoacidosis
C109C12	Type 2 diabetes mellitus with nephropathy	C10EN11	Type I diabetes mellitus with ketoacidotic coma
C10F200	Type 2 diabetes mellitus with neurological complications	C108B11	Type I diabetes mellitus with mononeuropathy
C109212	Type 2 diabetes mellitus with neurological complications	C108D11	Type I diabetes mellitus with nephropathy
C109H12	Type 2 diabetes mellitus with neuropathic arthropathy	C108211	Type I diabetes mellitus with neurological complications
C10FH00	Type 2 diabetes mellitus with neuropathic arthropathy	C108C11	Type I diabetes mellitus with polyneuropathy
C10F100	Type 2 diabetes mellitus with ophthalmic complications	C108011	Type I diabetes mellitus with renal complications
C109112	Type 2 diabetes mellitus with ophthalmic complications	C108711	Type I diabetes mellitus with retinopathy
C109F12	Type 2 diabetes mellitus with peripheral angiopathy	C108511	Type I diabetes mellitus with ulcer
C10FF00	Type 2 diabetes mellitus with peripheral angiopathy	C10EA11	Type I diabetes mellitus without complication
C10FM00	Type 2 diabetes mellitus with persistent microalbuminuria	C10F.11	Type II diabetes mellitus
C10FL00	Type 2 diabetes mellitus with persistent proteinuria	C109.13	Type II diabetes mellitus
C10FB00	Type 2 diabetes mellitus with polyneuropathy	C10F711	Type II diabetes mellitus - poor control
C109012	Type 2 diabetes mellitus with renal complications	C109711	Type II diabetes mellitus - poor control
C10F000	Type 2 diabetes mellitus with renal complications	C109G11	Type II diabetes mellitus with arthropathy

Code	Read/OXMIS name	Code	Read/OXMIS name
C109612	Type 2 diabetes mellitus with retinopathy	C109E11	Type II diabetes mellitus with diabetic cataract
C10F600	Type 2 diabetes mellitus with retinopathy	C109511	Type II diabetes mellitus with gangrene
C10F400	Type 2 diabetes mellitus with ulcer	C10F511	Type II diabetes mellitus with gangrene
C109D11	Type II diabetes mellitus with hypoglycaemic coma	C10E412	Unstable insulin dependent diabetes mellitus
C109A11	Type II diabetes mellitus with mononeuropathy	C108400	Unstable insulin dependent diabetes mellitus
C10F311	Type II diabetes mellitus with multiple complications	C10E400	Unstable type 1 diabetes mellitus
C109C11	Type II diabetes mellitus with nephropathy	C108411	Unstable type I diabetes mellitus
C10FC11	Type II diabetes mellitus with nephropathy	C10E411	Unstable type I diabetes mellitus
C109211	Type II diabetes mellitus with neurological complications	66AJ.11	Unstable diabetes
C109H11	Type II diabetes mellitus with neuropathic arthropathy	250 NT	UNSTABLE DIABETIC
C109111	Type II diabetes mellitus with ophthalmic complications	C10F611	Type II diabetes mellitus with retinopathy
C10FF11	Type II diabetes mellitus with peripheral angiopathy	C109411	Type II diabetes mellitus with ulcer
C109F11	Type II diabetes mellitus with peripheral angiopathy	C109911	Type II diabetes mellitus without complication
C10FL11	Type II diabetes mellitus with persistent proteinuria	C10F911	Type II diabetes mellitus without complication
C109B11	Type II diabetes mellitus with polyneuropathy	250 G	ULCER DIABETIC
C10FB11	Type II diabetes mellitus with polyneuropathy	9NND.00	Under care of diabetic foot screener
C109011	Type II diabetes mellitus with renal complications	9NN8.00	Under care of diabetologist
C10F011	Type II diabetes mellitus with renal complications	66A9.00	Understands diet - diabetes
C109611	Type II diabetes mellitus with retinopathy	C108z00	Unspecified diabetes mellitus with multiple complications

Appendix 10 – Read/OXMIS codes for biochemical ADRs

Code	Read/OXMIS name	Code	Read/OXMIS name
C367.00	Hyperkalaemia	C361.11	Hyponatraemia
7886K	HYPERKALAEMIA	7886N	HYPONATRAEMIA
C368.00	Hypokalaemia		
7887K	HYPOKALAEMIA		

Appendix 11 – Read/OXMIS for death codes

Code	Read/OXMIS name	Code	Read/OXMIS name
9234	FP22-death	9413	Med A given to family
94D..00	Hospital notified of death	T053z00	Killed by rolling stock - unspecified person
945..00	Hospital death discharge notif	T053000	Killed by rolling stock - railway employee
945Z.00	Hospital death disch. NOS	T140 FP	DEATH IN HOSPITAL
T0y0z00	Found dead on railway unspecified - unspecified person	T053300	Killed by rolling stock - pedal cyclist
T0y0200	Found dead on railway unspecified – pedestrian	T053y00	Killed by rolling stock - other specified person
9499	Found dead at accident site	22J3.00	O/E - dead - unattended death
22J..13	Died	ZV68011	[V]Issue of death certificate
8HG..00	Died in hospital	9433	Coroner report - paid for
T1400M	DIED	L0010GP	CORONER REFERRED TO
T0y0y00	Found dead on railway unspecified - other spec person	94B..11	Condition fatal-cause of death
T0y0.00	Found dead on railway right-of-way unspecified	941..11	Certificate - death
949..13	Died - place patient died	947..00	Cause of death clarif. SD17/18
T053200	Killed by rolling stock - pedestrian	94B..00	Cause of death
7962	FOUND DEAD	9452	Await hosp death disch letter
795 DR	DROPPED DEAD	9454	Ask for hosp death disch lett.
T0y0100	Found dead on railway unspecified - passenger	94...11	Administration after pat. died
22J4.00	O/E - dead - sudden death	TGyz400	Accidentally killed NOS
949B.00	Patient died in community hospital	RyuC100	[X]Other sudden death, cause unknown
9491	Patient died at home	946..00	Death notif.- non.hosp source
949..00	Patient died - to record place	RyuC.00	[X]Ill-defined and unknown causes of mortality
T400	PATIENT DIED	9441	Coroner's PM report awaited
22J..14	Patient died	R213z00	[D]Unattended death NOS
236..12	O/E - respiratory death	R213.00	[D]Unattended death
2329	O/E - death rattle	R21z.00	[D]Sudden death, cause unknown NOS
22JZ.00	O/E - dead NOS	R21..00	[D]Sudden death, cause unknown
22J1.00	O/E - dead - unexpected	R2...12	[D]Mortality, cause unsure
T053100	Killed by rolling stock - passenger	R211.00	[D]Instantaneous death
22J6.00	O/E - dead - suspicious death	R213100	[D]Found dead
T053.00	Killed by rolling stock	R213000	[D]Found after death, unknown cause of death
22J2.00	O/E - dead - expected	R212100	[D]Died, with no sign of disease
22J..11	O/E - dead - condition fatal	R212000	[D]Death, not instantaneous cause unknown
22J..00	O/E - dead	R212z00	[D]Death less than 24 hours from onset of illness NOS

Code	Read/OXMIS name	Code	Read/OXMIS name
9414	Med A not signed-coroner case	R212.00	[D]Death less than 24 hours from onset of illness
RyuC200	[X]Other ill-defined and unspecified causes of death	948..11	Stat B,C and F cremation certs
9498	Dead on arrival at hospital	G575100	Sudden cardiac death, so described
949C.00	Patient died in GP surgery	795 N	SUDDEN DEATH NONVIOLENT
9451	Death notif. from hospital	7963	UNKNOWN CAUSE DEATH
9492	Patient died in part 3 accom.	94A..00	Unexpected death-Coroner told
8HG..11	Death in hospital	9495	Patient died in hospital
941..00	Death certificate form Med A	949A.00	Patient died in hospice
9412	Death cert. Med A signed	9471	SD17/18 received-death clarif.
941Z.00	Death cert. Med A NOS	9494	Patient died in resid.inst.NOS
9411	Death cert. Med A due	T4001	VIOLENT DEATH
T140 FH	DEATH AT HOME	9493	Patient died in nursing home
9681D	DEATH ANAESTHETIC	T4002	SUDDEN DEATH
94Z..00	Death administration NOS	L0010GN	REFERRED TO CORONER
94...00	Death administration	949Z.00	Patient died in place NOS
9431	Coroner report - requested	9497	Patient died in publ.place NOS
22J..12	Death	9496	Patient died in street
9432	Coroner report - sent off	L 917PM	POST MORTEM REPORT RECEIVED
949..11	Dead - place patient died	949..14	Place of death
94E..00	Date of death	94A..11	Referral to coroner
948Z.00	Cremation certification NOS	94C..00	Post mortem report
948..00	Cremation certification	94C0.00	Post mortem report received
9484	Crem. form part C completed	9481	Patient for cremation
9483	Crem. form part C arranged	9453	Receiv hosp death disch letter
9482	Crem. form part B completed	7L1M000	Preoperative anaesthetic death
944..00	Coroner's post-mortem report	TK55.00	Suicide and selfinflicted injury by explosives
9442	Coroner's PM report requested	TKx3.00	Suicide and selfinflicted injury by extremes of cold
9443	Coroner's PM report received	TK5..00	Suicide and selfinflicted injury by firearms and explosives
944Z.00	Coroner's PM report NOS	TK30.00	Suicide and selfinflicted injury by hanging
949..12	Deceased - place patient died	TK50.00	Suicide and selfinflicted injury by handgun
T140 F	DEATH	TKx4.00	Suicide and selfinflicted injury by electrocution
947..12	SD18 - cause of death clarif	TKx5.00	Suicide and selfinflicted injury by crashing motor vehicle
943Z.00	Report for Coroner NOS	TK30.00	Suicide and selfinflicted injury by hanging
L 917WD	REPORT RECEIVED FROM CORONER	TK5z.00	Suicide and selfinflicted injury by firearms/explosives NOS
947..11	SD17 - cause of death clarif	TK4..00	Suicide and selfinflicted injury by drowning
947Z.00	SD17/18 cause of death NOS	TK6z.00	Suicide and selfinflicted injury by cutting and stabbing NOS
9472	SD17/18 completed	TK6..00	Suicide and selfinflicted injury by cutting and stabbing
943..00	Report for Coroner	TK60.00	Suicide and selfinflicted injury by cutting
9473	SD17/18-no details, returned	TKx6.00	Suicide and selfinflicted injury by crashing of aircraft
TKx1.00	Suicide and selfinflicted injury by burns or fire	TKx7.00	Suicide and selfinflicted injury caustic subst, excl poison

Code	Read/OXMIS name	Code	Read/OXMIS name
TKx1.00	Suicide and selfinflicted injury by burns or fire	TKx0100	Suicide + selfinflicted injury-lying before moving object
TK52.00	Suicide and selfinflicted injury by hunting rifle	TK05.00	Suicide + selfinflicted poisoning by drug or medicine NOS
TK21.00	Suicide and selfinflicted poisoning by other carbon monoxide	TK05.00	Suicide + selfinflicted poisoning by drug or medicine NOS
TK60.00	Suicide and selfinflicted injury by cutting	TK1z.00	Suicide + selfinflicted poisoning by domestic gases NOS
TK61.00	Suicide and selfinflicted injury by stabbing	TK07.00	Suicide + selfinflicted poisoning by corrosive/caustic subst
TK70.00	Suicide+selfinflicted injury-jump from residential premises	TK01.00	Suicide + selfinflicted poisoning by barbiturates
TK71.00	Suicide+selfinflicted injury-jump from oth manmade structure	TK01.00	Suicide + selfinflicted poisoning by barbiturates
TK72.00	Suicide+selfinflicted injury-jump from natural sites	TK08.00	Suicide + selfinflicted poisoning by arsenic + its compounds
TK7z.00	Suicide+selfinflicted injury-jump from high place NOS	TK00.00	Suicide + selfinflicted poisoning by analgesic/antipyretic
TK1y.00	Suicide and selfinflicted poisoning by other utility gas	TK0z.00	Suicide + selfinflicted poisoning by solid/liquid subst NOS
TK2y.00	Suicide + selfinflicted poisoning by other gases and vapours	TK06.00	Suicide + selfinflicted poisoning by agricultural chemical
TKz..00	Suicide and selfinflicted injury NOS	TK1..00	Suicide + selfinflicted poisoning by gases in domestic use
TK...00	Suicide and selfinflicted injury	TKx0000	Suicide + selfinflicted injury-jumping before moving object
TKz..00	Suicide and selfinflicted injury NOS	TKx0.00	Suicide + selfinflicted injury-jump/lie before moving object
TK61.00	Suicide and selfinflicted injury by stabbing	TK31.00	Suicide + selfinflicted injury by suffocation by plastic bag
U2...13	[X]Suicide	TK3..00	Suicide + selfinflicted injury by hang/strangulate/suffocate
TK51.00	Suicide and selfinflicted injury by shotgun	TKx0z00	Suicide + selfinflicted inj-jump/lie before moving obj NOS
TK51.00	Suicide and selfinflicted injury by shotgun	TK3y.00	Suicide + selfinflicted inj oth mean hang/strangle/suffocate
TKx2.00	Suicide and selfinflicted injury by scald	TK3y.00	Suicide + selfinflicted inj oth mean hang/strangle/suffocate
TKxy.00	Suicide and selfinflicted injury by other specified means	TK3z.00	Suicide + selfinflicted inj by hang/strangle/suffocate NOS
TKxz.00	Suicide and selfinflicted injury by other means NOS	3009D	SUICIDE
TKx..00	Suicide and selfinflicted injury by other means	TK00.00	Suicide + selfinflicted poisoning by analgesic/antipyretic
TK54.00	Suicide and selfinflicted injury by other firearm	TK0..00	Suicide + selfinflicted poisoning by solid/liquid substances
TK53.00	Suicide and selfinflicted injury by military firearms	TK01500	Suicide and self inflicted injury by Quinalbarbitone
TK01400	Suicide and self inflicted injury by Phenobarbitone	TK02.00	Suicide + selfinflicted poisoning by oth sedatives/hypnotics
TK01300	Suicide and self inflicted injury by Pentabarbitone	TK20.00	Suicide + selfinflicted poisoning by motor veh exhaust gas
TK01200	Suicide and self inflicted injury by Butabarbitone	TK11.00	Suicide + selfinflicted poisoning by liquified petrol gas
TK01z00	Suicide and self inflicted injury by barbiturates	TK...00	Suicide and selfinflicted injury
TK01100	Suicide and self inflicted injury by Barbitone	TK03.00	Suicide + selfinflicted poisoning tranquilliser/psychotropic

Code	Read/OXMIS name	Code	Read/OXMIS name
TK01000	Suicide and self inflicted injury by Amylobarbitone	TK0..00	Suicide + selfinflicted poisoning by solid/liquid substances
TK...14	Suicide and self harm	TK2..00	Suicide + selfinflicted poisoning by other gases and vapours
TK...14	Suicide and self harm	TK04.00	Suicide + selfinflicted poisoning by other drugs/medicines
TK10.00	Suicide + selfinflicted poisoning by gas via pipeline	TK04.00	Suicide + selfinflicted poisoning by other drugs/medicines
TK03.00	Suicide + selfinflicted poisoning tranquilliser/psychotropic	TK02.00	Suicide + selfinflicted poisoning by oth sedatives/hypnotics
TK2z.00	Suicide + selfinflicted poisoning by gases and vapours NOS		

Appendix 12 – Read/OXMIS codes for hospital admission

Code	Read/OXMIS name	Code	Read/OXMIS name
8Hd..00	Admission to hospital	ZLF2.00	Discharge from hospital
8H2Q.00	Admit cardiology emergency	8HE..00	Discharged from hospital
8H2O.00	Admit cardiothoracic emergency	8HE2.00	Discharged from inpatient care
8H2R.00	Admit COPD emergency	8HE4.00	Discharged from private hosp'l
8H2D.00	Admit dermatology emergency	8H23000	Emerg psychiatric admiss MHA
8H2J.00	Admit diabetic emergency	T927 E	EMERGENCY ADMISSION
8H29.00	Admit ENT emergency	8H2P.00	Emergency admission, asthma
8H24.00	Admit geriatric emergency	8H2..00	Emergency hospital admission
8H26.00	Admit gynaecological emergency	T927	HOSPITAL ADMISSION
8H2H.00	Admit haematology emergency	Y1190AC	HOSPITAL ADMISSION MENTAL HEALTH ACT
8H2S.00	Admit heart failure emergency	9b0K.00	Hospital admission note
8H2Z.00	Admit hospital emergency NOS	T932	HOSPITAL DISCHARGE
8H21.00	Admit medical emergency unsp.	13F8.11	Hospital inpatient
8H2E.00	Admit neurology emergency	9b0L.00	Hospital inpatient report
8H2N.00	Admit neurosurgical emergency	8CO..00	Inpatient care
8H27.00	Admit obstetric emergency	Z177800	Inpatient care
8H2B.00	Admit ophthalmological emerg.	T927 AA	INPATIENT HOSPITAL
8H2K.00	Admit oral surgical emergency	13F8100	Long stay hospital inpatient
8H28.00	Admit orthopaedic emergency	13FS.00	Long stay hospital inpatient
8H2I.00	Admit plastic surgery emergenc	T927 L	LONG-STAY HOSPITAL PATIENT
8H23.00	Admit psychiatric emergency	T306 HA	MEDICAL CERT FOR HOSPITAL ADMISSION
8H2L.00	Admit psychogeriatric emergency	8H3V.00	Non-urgent cardiological admission
8H2G.00	Admit radiotherapy emergency	8H3T.00	Non-urgent cardiothoracic admission
8H2M.00	Admit renal medicine emergency	8H3I.00	Non-urgent dermatology admisn.
8H2C.00	Admit rheumatology emergency	8H3O.00	Non-urgent diabetic admission
8H22.00	Admit surgical emergency unsp.	8H3E.00	Non-urgent ENT admission
8H15.00	Admit to burns unit	8H39.00	Non-urgent geriatric admission
8H11.00	Admit to cardiac ITU	8H3B.00	Non-urgent gynaecol.admission
8H1..11	Admit to I.T.U.	8H3M.00	Non-urgent haematology admisn.
8H1Z.00	Admit to intensive c.u. NOS	8H31.00	Non-urgent hosp.admission unsp
8H1..00	Admit to intensive care unit	8H3..00	Non-urgent hospital admission
8H14.00	Admit to metabolic ITU	8H36.00	Non-urgent medical admission
8H13.00	Admit to neurological ITU	8H3J.00	Non-urgent neurology admission
8H12.00	Admit to respiratory ITU	8H3S.00	Non-urgent neurosurgical admission
8H2A.00	Admit trauma emergency	8H3C.00	Non-urgent obstetric admission
8H2F.00	Admit urology emergency	8H3G.00	Non-urgent ophthalmolog.admisn
T927 MT	ADMITTED MENTAL HOSPITAL	8H3U.00	Non-urgent oral Surg.admission
T9270AC	COMPULSORY ADMISSION TO HOSPITAL	8H3D.00	Non-urgent orthopaedic admisn.
8H3D.00	Non-urgent orthopaedic admisn.	8HJ..00	Self-referral to hospital
8H3N.00	Non-urgent plastic surg.admisn	8BAR.00	Specialist palliative care treatment - inpatient
8H38.00	Non-urgent psychiatric admisn.	8HF..12	Transferred from hospital
8H3Q.00	Non-urgent psychogeriatric admission	9N1B.00	Seen in hospital ward
8H3L.00	Non-urgent radiotherapy admisn	8H3Z.00	Other hospital admission NOS
8H3R.00	Non-urgent renal medicine admission	9144	Patient in hospital
8H3P.00	Non-urgent respiratory admission	8H3K.00	Non-urgent urology admission
8H3H.00	Non-urgent rheumatology admisn	8H3F.00	Non-urgent trauma admission
8H37.00	Non-urgent surgical admission		

Appendix 13 – Stata® codes for propensity score matching

```
use "C:\Users\Sarah\Documents\PhD\Propensity score analysis -
baseline monitoring.dta", clear
**** Run psmatch2 based on pscore1009 ****
gen u = uniform()
sort u
psmatch2  anybaselinetesting6mos, pscore(pscore1009) caliper (0.001)
descending noreplacement common
gen casecontrolpscore1009 = 1 if _treated==1 & _weight==1
replace casecontrolpscore1009 = 0 if _treated==0 & _weight==1
tab casecontrolpscore1009

*** Create matching ID for conditional LR analysis ***
gen matchpscore1009 = _id if _treated==0 & _weight==1
replace matchpscore1009 = _n1 if _treated==1 & _weight==1

*** Assess balance on variables ***
pstest ageatevdate
pstest Diabetesstatuscorrected
pstest evyear2000
pstest evyear2001
pstest evyear2002
pstest evyear2003
pstest Smokingbinary
pstest gender
pstest SESscore
pstest BMI
pstest unitsweek
pstest SystolicBP
pstest DiastolicBP
pstest hypothyroidism
pstest PVD
pstest carbamazepinebaseline
pstest nsaidbaseline
pstest SSRIbaseline
pstest tricyclic
pstest trimethoprimbaseline
pstest aceibaseline
pstest alphabaseline
pstest at2baseline
pstest betabaseline
pstest cabaseline
pstest loopbaseline
pstest mixedbaseline
pstest thiazidebaseline
pstest gpvisits8plus
pstest rx1year10plus
pstest elecdisturb


*** compare 2 groups in matched cohort ***
ttest ageatevdate, by(casecontrolpscore1009)
tab gender casecontrolpscore1009, chi column
ttest SESscore, by(casecontrolpscore1009)
ttest BMI, by(casecontrolpscore1009)
ttest unitsweek , by(casecontrolpscore1009)
ttest SystolicBP, by(casecontrolpscore1009)
ttest DiastolicBP, by(casecontrolpscore1009)
tab Diabetesstatuscorrected casecontrolpscore1009, chi column
```

```
tab Smokingbinary  casecontrolpscore1009, chi column
tab hypothyroidism   casecontrolpscore1009, chi column
tab PVD  casecontrolpscore1009, chi column
tab aceibaseline  casecontrolpscore1009, chi column
tab at2baseline   casecontrolpscore1009, chi column
tab betabaseline   casecontrolpscore1009, chi column
tab cabaseline   casecontrolpscore1009, chi column
tab mixedbaseline  casecontrolpscore1009, chi column
tab loopbaseline casecontrolpscore1009, chi column
tab thiazidebaseline  casecontrolpscore1009, chi column
tab carbamazepinebaseline  casecontrolpscore1009, chi column
tab nsaidbaseline   casecontrolpscore1009, chi column
tab SSRIbaseline   casecontrolpscore1009, chi column
tab tricyclic  casecontrolpscore1009, chi column
tab trimethoprimbaseline casecontrolpscore1009, chi column
tab evyear2000 casecontrolpscore1009, chi column
tab evyear2001 casecontrolpscore1009, chi column
tab evyear2002 casecontrolpscore1009, chi column
tab evyear2003 casecontrolpscore1009, chi column
tab rx1year10plus  casecontrolpscore1009, chi column
tab elecdisturb  casecontrolpscore1009, chi column
tab gpvisits8plus   casecontrolpscore1009, chi column

*** Q-Q plots ***
qqplot  pscore1009base pscore1009nobase if pscore1009q==1
qqplot  pscore1009base pscore1009nobase if pscore1009q==2
qqplot  pscore1009base pscore1009nobase if pscore1009q==3
qqplot  pscore1009base pscore1009nobase if pscore1009q==4
qqplot  pscore1009base pscore1009nobase if pscore1009q==5


** Boxplot of quintiles ***
graph box pscore1009 if drugsonevdate==1,
//over(anybaselinetesting6mos) over(pscore1009q)

*** Draw kensity plots to demonstrate whether matching has been
successful ***
kdensity pscore1009 if anybaselinetesting6mos==1, addplot((kdensity
//pscore1009 if anybaselinetesting6mos==0))
kdensity pscore1009 if casecontrolpscore6==1, addplot((kdensity
//pscore1009 if casecontrolpscore6==0))

*** Run clogit with matchpscore1009 as a matching id ***

clogit HospAdmCode6mos anybaselinetesting6mos, group(matchpscore1009)
or
clogit AnyADRCode6mos anybaselinetesting6mos, group(matchpscore1009)
or
clogit DeathCode6mos anybaselinetesting6mos, group(matchpscore1009)
or
```

Work related to this thesis

Peer-reviewed publications

McDowell, S. E., Coleman, J. J., Evans, S. J. W., Gill, P. S., Ferner, R. E. (2010) Laboratory monitoring and adverse patient outcomes with antihypertensive therapy in primary care. **Pharmacoepidemiology and Drug Safety,** 19 (5): 482–489.

Coleman, J. J., **McDowell, S. E.**, Evans, S. J. W., Gill, P. S., Ferner, R. E. (2010) Oversight: a retrospective study of biochemical monitoring in patients beginning antihypertensive drug treatment in primary care. **British Journal of Clinical Pharmacology,** 70 (1): 109–117.

McDowell, S. E., Ferner R. E. Biochemical monitoring of patients treated with antihypertensive therapy for adverse drug reactions: a systematic review (accepted for publication in Drug Safety).

Other publications

McDowell, S. E. (2010) Monitoring of patients treated with antihypertensive therapy for adverse drug reactions. **Adverse Drug Reaction Bulletin**, 261: 1003–1006.

Oral presentations at conferences

McDowell, S. E*., Coleman, J. J., Evans, S. J. W., Ferner, R. E. The relationship between baseline testing of serum electrolytes and creatinine and adverse outcomes in patients treated with antihypertensive drugs: an analysis using propensity score matching methods. Presented at the 9th Annual Meeting of the International Society of Pharmacovigilance, 7–9 October 2009, Reims, France. **Speaker*

Poster presentations at conferences

McDowell, S. E., Coleman, J. J., Evans, S. J. W., Ferner, R. E. Serum sodium concentration monitoring and hyponatraemia in thiazide-treated patients. Presented at the 26th International Conference on Pharmacoepidemiology & Therapeutic Risk Management, 19–22 August 2010, Brighton, UK.

Coleman, J. J., **McDowell, S. E.**, Evans, S. J. W., Gill, P. S., Ferner, R. E. The monitoring of serum electrolytes and creatinine in patients treated with antihypertensive drugs: an analysis of the General Practice Research Database. Presented at the 8th Annual Meeting of the International Society of Pharmacovigilance, 5–8 October 2008, Buenos Aires, Argentina.

Coleman, J. J., **McDowell, S. E.**, Ferner, R. E. Where does the evidence for guidelines come from? An example of monitoring for the adverse renal effects of ACE inhibitors. Presented at the 8th Annual Meeting of the International Society of Pharmacovigilance, 5–8 October 2008, Buenos Aries, Argentina.

Presentations at professional meetings

McDowell, S.E. Dealing with missing data in observational studies: the GPRD as an example. Presented at the Clinical Pharmacology Colloquium, Liverpool, UK, 15 May 2010.

McDowell, S.E. Propensity score methods in large observational databases: the GPRD as an example. Presented at the Clinical Pharmacology Colloquium, Oxford, UK, 20 June 2009.

McDowell, S.E., Coleman JJ, Evans SJW, Ferner RE. Monitoring for adverse drug reactions in the General Practice Research Database. Presented at the Clinical Pharmacology Colloquium, Warwick, UK, 15 November 2008.