THE NATURAL HISTORY OF IMPAIRED GLUCOSE REGULATION AMONGST YOUNG ADULTS AND ITS EFFECT ON RENAL FUNCTION

---

by

FEROZKHAN JADHAKHAN

A thesis submitted to the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Primary Care Clinical Sciences

Institute of Applied Health Research

College of Medical and Dental Sciences

University of Birmingham

November 2016

Abstract

Introduction: The impact of diabetes as a major cause of chronic kidney disease (CKD) is established however the risk associated with impaired glucose regulation (IGR) is not well characterised. It is not clear whether individuals with IGR are at risk of developing CKD or whether the risk is confined to individuals who progress to overt diabetes.

Objective: To determine the risk of CKD in young adults aged 18 to 40 years with IGR compared to those with normoglycaemia.

Methods: This study consists of three main parts: 1) systematic review, 2) analysis to determine incidence and period prevalence of IGR, and 3) analysis to determine risk of CKD in IGR. A systematic review was undertaken to estimate the risk of CKD associated with IGR. MEDLINE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), EMBASE, PubMed, Database of Abstracts of Reviews of Effects (DARE), Cochrane Database of Systematic Reviews (CDSR) and Trip Database were searched for cohort and case control studies comparing risk of CKD in individuals with and without IGR. A retrospective cohort study was undertaken using a large dataset of patient records to describe the incidence and prevalence of IGR and to investigate the relationship between IGR and CKD. Patients diagnosed with IGR (2000 - 2015) registered at a practice from across the UK contributing data to The Health Improvement Network (THIN) database were included in the analyses. Read coded diagnoses and clinical measurements were used to ascertain incident CKD and selected covariates. Incidence rate of IGR per 100,000 person-years and period prevalence were reported by age group, sex, ethnicity, area of deprivation and calendar year. Poisson regression model were used to obtain adjusted incident rate ratio of CKD. Furthermore, incidence rate of CKD was calculated by CKD category (stage 1 – 2 /3 – 5) in individuals with IGR. Cox proportional hazards model was fitted to estimate the risk of CKD

following IGR diagnosis. The five year survival probability of remaining free of CKD was estimated using Kaplan-Meier survivor function.

Results: The systematic review found no evidence associating risk of CKD in young adults aged 18 to 40 years with IGR. Sufficient studies were not available for a meta-analysis hence quantification of CKD risk was not possible. The THIN database analysis shows that incidence of recorded IGR increased 8.4 times between 2000 and year 2014. Recorded incidence was higher in those aged 26 to 40 years than those aged 18 to 25 years, in women than men, in more deprived than less deprived areas and in Black and South Asian than White ethnic groups. Prevalence was significantly higher for those aged 26 to 40 years than those aged 18 to 25 years. Prevalence increased consistently across age groups for the period 2000 - 2014. Prevalence in females was consistently higher than males throughout the 14 year period. During 2000 – 2014, prevalence of diagnosed IGR increased among both males and females. Incidence of CKD was 4 times higher in IGR than normoglycaemia, after adjusting for confounders risk of CKD was reduced to 2.6 times. The incidence of CKD stage (3 – 5) was approximately 4 times higher than the incidence of CKD stage (1 – 2) in the IGR cohort. The predictors of CKD in IGR model shows that patients with hypertension were approximately 3 times more at risk of CKD and for each additional year of age at IGR diagnosis, CKD risk increased by 7%. The overall proportion of variation explained by the model was 24% (R^2 0.24). The Kaplan-Meier (K-M) estimates show that 2% of IGR patients were diagnosed with CKD by 2 years increasing to 4% after 5 years follow-up.

Conclusion: Results of the systematic review demonstrate that the risk of CKD in young adults with IGR remains to be elucidated, as no evidence was found associating risk of CKD in young adults with IGR. The THIN database analyses provide evidence of an increased risk of CKD amongst young adults with IGR. It also showed that IGR patients are at higher risk

of CKD stage (3 – 5) compared to CKD stage (1 – 2). Among the modifiable risk factors, hypertension was consistently linked to higher incidence of CKD.

Dedication

I affectionately dedicate this thesis to my family for their unconditional love, support and guidance.

Acknowledgements

The completion of this PhD has been a long and at times an arduous journey which would not have been possible without the unconditional support, guidance and invaluable insight that I have received from many people. First and foremost, I would like to express my deepest gratitude to my supervisors, Doctor Paramjit Gill, Professor Tom Marshall and Dr Ronan Ryan for their continuous support, motivation, enthusiasm, constant feedback and above all, their extensive knowledge and illuminating view which has helped me through not only writing of this PhD but through some of the hardest times of my life. A special thank you to Doctor Ronan Ryan, whose wisdom, incredible patience, vision and constructive feedback has helped me navigate safely through the THIN database analysis.

I wish to express my sincere gratitude to the National Institute for Health Research (NIHR) School for Primary Care Research (SPCR) for funding this PhD and for providing excellent training opportunities.

I am deeply grateful to my wife Anya for her tireless support and understanding of my aspirations and goals. It hasn't been easy at times but you showed your infallible love and support throughout this PhD and my life. Without your help I wouldn't have been where I am today. A special mention of my 4 months old son Isaac and daughter Mia love you both. Heartfelt thanks to my dear friend Sonal who agreed to proofread this thesis.

Particular thanks go to my parents who have been and remained central in my life, how can I ever repay you for the sacrifice you have made over the years and for giving me the upbringing that made me the person I am today. I am ever so grateful for your unflinching love, support and generosity. A heartfelt thanks to my little sister Noorina and younger brother Naheed, you have always been there for me. Although great distance separates us, we are spiritually connected.

Statement of contribution

All the chapters within this thesis were composed by myself and continuously reviewed by my supervisors, Professor Paramjit Gill, Professor Tom Marshall and Doctor Ronan Ryan. The concept, design, method, analysis and interpretation of data were extensively examined by my supervisors. Professor Tom Marshall and Doctor Paramjit Gill were involved in the design and concept of the systematic review protocol and the full review. They provided feedback on the published protocol (Appendix 1) and were the second reviewers for the full review (Appendix 3 and Chapter 3). Doctor Ronan Ryan provided advice and guidance for the design, analysis and interpretation of results (chapters, 6 and 7) for the THIN database studies including the prognostic study and extracted data for chapters (4 and 5).

List of publications

- **Jadhakhan F**, Marshall T and Gill P. A systematic review investigating the cumulative incidence of chronic kidney disease in young adults with impaired glucose tolerance. Sys Rev 2015; 4(69): 1-6.
- **Jadhakhan F**, Marshall T and Gill P. A systematic review exploring the effects of impaired glucose tolerance (IGT) on incidence of chronic kidney disease (CKD) in young adults. Br J Diabet 2016; 16(4): 162-67.
- Leong WB, **Jadhakhan F**, Taheri S, Chen YF, Adab P and Thomas GN. Effect of obstructive sleep apnoea on diabetic retinopathy and maculopathy: a systematic review and meta-analysis. Diabet Med 2016; 33(2): 158 – 68.
- Leong WB, **Jadhakhan F**, Taheri S, Thomas GN and Adab P. The Association between Obstructive Sleep Apnoea on Diabetic Kidney Disease: A Systematic Review and Meta-Analysis. Sleep 2016; 39(2): 301-8

List of tables

List of figures

Lists of abbreviations

Abbreviation		Terms
A	**AKI**	Acute Kidney Injury
	ACR	Albumin Creatinine Ratio
	ADA	American Diabetes Association
	AHD	Additional Health Data
	AMR	Acceptable Mortality Recording
	AF	Atrial Fibrillation
	AUS-DIAB	Australian Diabetes, Obesity and Lifestyle Study
	ARIC	Atherosclerosis Risk in Community Study
B	**BMI**	Body Mass Index
	BNF	British National Formulary
	BP	Blood Pressure
C	**CKD**	Chronic Kidney Disease
	Cr-EDTA	Chromium-51 Ethylenedamine Tetraacetic Acid
	CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
	CVD	Cardiovascular Disease
	CHD	Coronary Heart Disease
	CPRD	Clinical Practice Research Datalink
	CrCl	Creatinine Clearance
	CINAHL	Cumulative Index to Nursing and Allied Health Literature
	CDSR	Cochrane Database of Systematic Reviews
	CPCI	Conference Proceedings Citation Index
	CURES	Chennai Urban-Rural Epidemiology Study
	C-G	Cockroft Gault
	CI	Confidence Interval
D	**DTPA**	Diethylenetriamine Pentaacetic Acid
	DIN	Doctors Independent Network
	DBP	Diastolic Blood Pressure
	DECODE	Diabetes Epidemiology: Collaborative Analysis of Diagnostic Criteria in Europe
E	**EGFR**	Estimated Glomerular Filtration Rate
	ESRD	End Stage Renal Disease
	EMIS	Egton Medical Information Systems

Abbreviation continued on next page

Abbreviation continued from next page

Abbreviation		Terms
F	**FPG**	Fasting Plasma Glucose
G	**GPRD**	General Practice Research Database
	GP(s)	General Practitioner(s)
	GMS	General Medical Services
H	**HTA**	Health Technology Assessment
	HUNT II	Health Survey of Nord-Trøndelag County
	HR	Hazards Ratio
	HbA1c	Glycosylated Haemoglobin
	HSE	Health Survey for England
	HES	Hospital Episodes Statistics
	HSCIC	Health and Social Care Information Centre
	HF	Heart Failure
	HDL	High Density Lipoprotein
I	**IGT**	Impaired Glucose Tolerance
	IFG	Impaired Fasting Glucose
	IEC	International Expert Committee
	IDF	International Diabetes Federation
	IGR	Impaired Glucose Regulation
	InPs	In Practice Systems
	IRR	Incidence Rate Ratio
	IR	Incidence Rate
	IQR	Interquartile Range
K	**KDIGO**	Kidney Disease Improving Global Kidney Outcomes
	K/DOGI	Kidney Disease Outcomes Quality Initiative
	K-M	Kaplan Meier
M	**MDRD**	Modification of Diet in Renal Disease
	MHRA	Medicines and Healthcare Products Regulatory Agency
	MSGP	Morbidity Statistics from General Practice
	MOOSE	Meta-Analysis of Observational Studies in Epidemiology Group
	MI	Myocardial Infarction
N	**NICE**	National Institute for Healthcare Excellence
	NKF	National Kidney Foundation
	NKDEP	National Kidney Disease Education Program
	NEOERIC	New Opportunities for Early Renal Intervention by Computerised Assessment
	NHANES	National Health and Nutrition Examination Survey

Abbreviation continued on next page

Abbreviation continued from next page

Abbreviation		Terms
	NGT	Normal Glucose Tolerance
	NIDDM	Non-Insulin Dependent Diabetes Mellitus
	NOS	Newcastle-Ottawa Scale
	NSAIDS	Non-Steroidal Anti-Inflammatory Drugs
	NHS-DPP	National Health Service: Diabetic Prevention Programme
	NCEP	National Cholesterol Education Program
	NR	Not Reported
	NIHR	National Institute for Health and Care Excellence
O	**OR**	Odds Ratio
	OGTT	Oral Glucose Tolerance Test
	ONS	Office for National Statistics
P	**PPV**	Positive Predictive Value
	PCR	Protein Creatinine Ratio
	PRISMA	Preferred reporting Items for Systematic Reviews and Meta-Analysis
	PYR	Person Years
	PROSPERO	Prospective Registering of Systematic Reviews
Q	**QOF**	Quality and Outcomes Framework
R	**RR**	Relative Risk
S	**SCR**	Serum Creatinine
	SCI	Science Citation Index
	SBP	Systolic Blood Pressure
	STAND	Sedentary Time And Diabetes
T	**T2DM**	Type 2 Diabetes
	THIN	The Health Improvement Network Database
	τ^2	Tau squared
U	**UK**	United Kingdom
	UKPDS	United Kingdom Prospective Diabetic Study
V	**VAMP**	Value Added Medical Products
W	**WHO**	World Health Organisation
X	χ^2	Chi-square

TABLE OF CONTENTS

CHAPTER ONE
GENERAL INTRODUCTION
BACKGROUND, PREVALENCE, AETIOLOGY OF CHRONIC KIDNEY DISEASE (CKD) AND IMPAIRED GLUCOSE REGULATION (IGR)

1.1 Introduction to chapter

Chronic kidney disease (CKD) is defined as a gradual loss of kidney function for more than three months with or without kidney damage (1). It is a continuum of kidney conditions encompassing mild kidney damage to the most serious form of end stage renal disease (ESRD). CKD is now recognised as a major public health threat resulting in an increase in mortality, morbidity and poor quality of life (2). It is common, progressive, usually asymptomatic (3) and can co-exist with other conditions (3).

Impaired glucose regulation (IGR) is an intermediate state between normal glucose homeostasis and diabetic hyperglycaemia. Individuals with IGR have glucose levels higher than normal but not high enough for a diagnosis of diabetes (4).

IGR has been shown to be linked with increased risk of cardiovascular outcomes. Although the risk of CKD in patients with diabetes is relatively well established, the risk of young adults aged 18 to 40 years with IGR developing CKD is not well characterised. There is some evidence that the incidence of CKD is elevated in individuals with IGR, but this is confined to specific populations (5-7). Furthermore, progression of IGR to type 2 diabetes (T2DM) and subsequent CKD development is not well understood. Research is clearly warranted to establish a reliable estimate of the incidence of CKD in young adults aged 18 to 40 years with IGR.

1.2 Heterogeneous nature of CKD

CKD is a general term frequently used to describe numerous renal disorders and processes, some of which are better defined than others. This variation is explained to a certain extent by the cause, severity and rate of progression of the disease. The term was coined to simplify recognition and classification of the disease. Treatment and management of these unique conditions will significantly differ in many important aspects. Subsequently, Kidney Disease Improving Global Kidney Outcomes (KDIGO), an international independent non-profit organisation whose mission is to improve the care and outcomes of CKD patients provided important guidelines on CKD classification and staging in recognition to these important differences (8). This is further discussed in the next section.

1.3 Definitions, staging and measurement of CKD

1.3.1 Definition and staging of CKD

Chronic kidney disease has been defined as an estimated glomerular filtration rate (eGFR) of <60ml/min/1.73m^2 (standard body size) for at least 3 months or pathological abnormalities in the composition of blood, urine or abnormalities in imaging test for ≥3 months with or without reduced kidney function (1). Furthermore, the National Institute for Health and Care Excellence (NICE) recommend that an eGFR of (≤60ml/min/1.73m^2) along with a serum creatinine reduction of >20% should be regarded as clinically significant (9). The National Kidney Foundation (NKF) Kidney Disease Outcomes Quality Initiative (K/DOQI) in 2002 introduced clinical practical guidelines on the evaluation, classification and risk stratification of CKD (10). This was endorsed by KDIGO in 2004 (11). However, in 2009 KDIGO initiated and sponsored a Controversies Conference to examine the current definition and classification of CKD. It was acknowledged that the (K/DOQI) definition and staging was deficient because it was largely based on cross-sectional data and association of adverse clinical outcomes with levels of eGFR, albuminuria and proteinuria were poorly reported.

Furthermore, multiple meta-analyses of cohort studies have reported eGFR levels to be independently associated with cardiovascular mortality, ESRD and acute kidney injury (AKI) (12-14). The conference report recommended that both GFR and albuminuria should be used to define and classify CKD for a better estimation of CKD prognosis in patients (15). The thresholds of GFR categories remained largely the same but CKD category 3 was sub-divided into categories 3a (<60 ml/min/1.73m^2) and 3b (<45ml/min/1.73m^2) which represent mildly to moderately decreased and moderately to severely decreased kidney function. The updated guideline retained the original definition of GFR (<60ml/min/1.73m^2) but added albumin to creatinine ratio (ACR) of (≥30mg/mmol) as criteria for CKD (Table 1).

Table 1: Classification of chronic kidney disease using GFR and ACR categories:

GFR and ACR categories and risk of adverse outcomes			ACR categories (mg/mmol), description and range		
			<30 Normal to mildly increased	3-30 Moderately increased	>30 Severely increased
			A1	A2	A3
	≥90 Normal to high	G1			
	60-89 Mild reduction related to normal range for a young adult	G2			
	45-59 Mild-moderate reduction	G3a			
	30-44 Moderate – severe reduction	G3b			
	15-29 Severe reduction	G4			
	<15 Kidney failure	G5			

Increased risk of progressive CKD

Increased risk of progressive CKD

Adapted from Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group (2013) KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease, Kidney International (Suppl.3): 1-150

1.3.2 Renal function assessment

1.3.2.1 Glomerular filtration rate (GFR)

The glomerular filtration rate is the volume of blood filtered through by the kidney's glomeruli per unit time, measured in (ml/min). It is a measure of how well the kidneys are filtering waste products and maintaining fluid balance. GFR is the commonly accepted overall standard measure of renal function. It is usually detected by measurement of both exogenous and endogenous filtration markers (16). Exogenous measurements include Inulin, iohexol, ^{51}Cr-EDTA, ^{125}I-iothalamate and ^{99m}Tc-diethylenetriaminepentaacetic acid (DTPA). These are considered as gold standard in yielding reliable results. These tests however are expensive and time consuming and can only be performed in specialised laboratories. Endogenous markers (creatinine and cystatin C) are much more commonly used to estimate GFR. They are cheaper and more simple but less reliable (17). Level of GFR fluctuates with age, sex, body size and haemodynamic factors (18). Calculating GFR is usually complex and expensive, making it unsuitable for routine assessment of kidney function. Instead, estimated glomerular filtration rate (eGFR) calculated using a prediction equation is reported by laboratories as a practical test of kidney function (9). The most widely used equation is the abbreviated Modification of Diet in Renal Disease (MDRD); it is calculated from serum creatinine, sex, age and race (19). On the other hand, Cockcroft-Gault was developed to estimate creatinine clearance. The equation is based on serum creatinine, age, gender and weight (20). Prediction equations such as MDRD and Cockcroft-Gault are frequently used to ascertain eGFR based on creatinine and cystatin C.

1.3.2.2 Serum creatinine

Serum creatinine level is the most widely used parameter to assess renal function in clinical practice. During muscle metabolism creatinine is produced as a by-product and for the majority of people remains stable as muscle mass varies little from day to day. Creatinine is a

low molecular weight cation which is distributed throughout the body and freely filtered by the renal glomerulus (21). The serum creatinine level depends on the glomerular filtration rate (GFR), during renal dysfunction the creatinine filtration is reduced and serum creatinine rises. This has been found to be a useful and accurate enough marker to evaluate renal function (22). Unfortunately, serum creatinine is also known to be a poor marker of GFR and insensitive to even mild to moderate decrease in GFR (23). Furthermore, it does not take into account variations in demographics, ethnicity, body mass, diet and medications which all affect creatinine concentration and therefore GFR (24).

1.3.2.3 Cystatin C

Cystatin C is a small protein produced by all nucleated cells at a steady rate. It is freely filtered by the glomerulus and almost completely reabsorbed by the renal tubular cells. Also its production does not appear to be influenced by gender, body mass and diet (25). Multiple studies have confirmed cystatin C as a superior marker in estimating GFR than serum creatinine (26-28). Additionally, cystatin C has been shown to be particularly sensitive in patients with mild to moderate changes in GFR (29). Furthermore, the accuracy of GFR estimating equations (MDRD and Chronic Kidney Disease Epidemiology Collaboration CKD-EPI) incorporating cystatin C and adjusting for ethnicity, proteinuria and diabetes is being investigated in a novel health technology assessment (HTA) funded prospective longitudinal study in a multi-ethnic population with stage 3 CKD across six UK centres (30). These equations are further discussed below.

1.3.2.4 Proteinuria and Albuminuria

Proteinuria is the most common and sensitive marker of progressive renal disease (31). Furthermore, albuminuria has been shown to be a predictor of renal impairment in the general population independent of age, gender and cardiovascular risk factors (32). Numerous studies in different population have evaluated the role of proteinuria and albuminuria in the

progression of renal disease. In a multi-centre randomised controlled trial to determine whether the risk of developing end stage renal disease was associated with multiple risk factors including proteinuria, a total of 12,866 men without kidney disease at baseline but at risk of heart disease were recruited between 1973 and 1975 and followed-up through 1999. Proteinuria was measured by urine dipstick and categorised as negative/trace, 1+ and ≥2+. The adjusted hazard ratio (HR) for developing ESRD in men with 1+ proteinuria was 2.30 (95% confidence interval [CI], 1.28 to 4.13) and 14.21 (95% CI, 9.16 to 22.05) in men with 2+ proteinuria (33). Additionally, in a study of 917 non-diabetic hypertensive patients free from renal disease at baseline and followed-up for 11.8 years, micro-albuminuria was found to be associated with an increased risk of developing chronic renal insufficiency with a relative risk (RR) of 7.61 (95% CI, 3.91 to 8.16) (34).

1.3.3 What are the available equations for estimating GFR?

1.3.3.1 Modification of Diet in Renal Disease (MDRD)

The MDRD equation was developed in 1999 from a study population of 1628 non-diabetic predominantly Caucasian men and women, aged 18 to 70 years with CKD. The equation was based on 6 variables namely: age, sex, serum creatinine, urea, albumin and ethnicity (35). A simplified version of the MDRD equation was introduced in 2000, the abbreviated 4-variables (age, sex, serum creatinine and ethnicity) was demonstrated to have greater precision and accuracy in predicting GFR (Figure 1) (36).

Figure 1: MDRD 4-variables equation

eGFR (mL × min–1 × [1.73 m2] –1) = 175 × (SCr standardized [mg × dL–1]) –1.154 × (age [years]) – 0.203

Gender correction:

Women	× 0.742
Blacks	×1.18

1.3.3.2 Cockroft – Gault (C-G) formula

The C-G formula was originally developed to predict creatinine clearance from serum creatinine, age and weight. The formula (Figure 2) was subsequently re-expressed to estimate GFR (20).

Figure 2: The Cockroft-Gault equation to estimate GFR

$$\frac{[140 - \text{AGE}] \times \text{WEIGHT (KG)}}{\text{PCr} \times 72 \times 0.85}$$

Gender correction:

Female: 0.85
Male: 1.00

1.3.3.3 Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI)

The chronic disease collaboration (CKD-EPI) was published in 2009 and development of the equation included participants with or without kidney disease from various clinical settings with diverse clinical characteristics and a wide range of GFR measurements. The equation uses the same variables (serum creatinine level, age, sex, and race) as the MDRD model. The (CKD-EPI) model (Figure 3) has been shown to have greater precision in estimating GFR than the MDRD equation especially in patients with higher GFR values (37).

Figure 3: The CKD-EPI equation

$$GFR = 141 \times \min(SCr/k, 1)^{\alpha} \times \max(SCr/k, 1)^{-1.209} \times 0.993^{Age} \times 1.018\ [\text{if female}] \times 1.159\ [\text{if black}]$$

Where:
SCr is serum creatinine in mg/dL
k is 0.7 for females and 0.9 for males
α is -0.329 for females and -0.411 for males
min indicates the minimum of SCr/k or 1, and
max indicates the maximum of SCr/k or 1

1.4 Advances in estimating GFR

Predictive equations (MDRD and C-G) provide a rapid method of estimating GFR and assessing renal function in patients with kidney disease. These formulas however are limited

by a lack of validation in the full range of GFR and ethnic diversity to which they are applied. Both equations greatly overestimate the strength of association between GFR and serum creatinine. These equations were reasonably accurate in estimating GFR in patients with moderate to advanced CKD but were poor in estimating GFR in healthy population. This could mean that underestimation of GFR could potentially lead to individuals labelled as false positive (38). Variability in laboratories calibrating serum creatinine and subsequent estimation of GFR from serum creatinine based equations can potentially introduce errors in estimating GFR. To overcome these limitations, the National Kidney Disease Education Program (NKDEP) introduced usage of calibrated serum creatinine assays based on isotope dilution mass spectrometry method which on average leads to lower level of serum creatinine but higher values for eGFR (36). The accuracy of eGFR estimation is likely to further improve by the introduction of more accurate prediction models such as the Chronic Kidney Disease Epidemiology (CKD-EPI) (39) or alternative prediction markers such as cystatin-C in CKD staging (40) and application of novel biomarker such as kidney injury molecule- 1 and clusterin (41).

1.5 Prevalence of CKD: United Kingdom and International Perspectives

1.5.1 UK prevalence of CKD

A number of studies have estimated the prevalence of CKD in the UK. The Health Survey for England showed that the overall prevalence of CKD stage 3-5 based on eGFR (<60ml/min/1.73m2) was 6% in men and 7% in women. Serum creatinine levels were used to estimate eGFR, using the MDRD equation. The survey showed wide variation by age, with 1% in men and 2% in women aged 16 to 54 years at stage 3-5. Prevalence rose significantly to 31% of men and 36% women aged 75 and over. The presence of albuminuria was found in 9% of men and 8% of women. In most cases micro-albuminuria was most prevalent with 8% in both men and women (42).

The NEw Opportunities for Early Renal Intervention by Computerised Assessment (NEOERICA) project, a large UK primary care study (practice population: 162,113) showed an age standardised prevalence of CKD stage 3-5 of 8.5% (10.6% females and 5.8% males). The prevalence of CKD and age shows a striking relationship, with approximately 70% of patients diagnosed with CKD aged (≥65 years) (43). CKD was determined by serum creatinine measured over a period of 1 year and eGFR calculated by MDRD equation. Additionally, the Quality and Outcomes Framework (QOF) data for England showed that the prevalence of diagnosed CKD (eGFR<60ml/min/1.73m^2) in patients aged (≥18 years) was 4.3%. The QOF data shows significant variation in CKD prevalence between health authorities, ranging from 1.4% to 9.6% in 2010/11 (9). Furthermore, the QOF register provided separate CKD stage 3-5 estimates defined by an eGFR (<60ml/min/1.73m^2) in individuals aged (≥18 years) for England, Scotland, Wales and Northern Ireland (Table 2).

Table 2: CKD prevalence in the UK

Country	CKD prevalence (%)
England (44)	4.3
Scotland(45)	3.23
Wales (46)	3.6
Northern Ireland(47)	5.0

1.5.2 Global prevalence of CKD

Worldwide, approximately 10% of the adult population is affected by some form of CKD and millions die prematurely each year due to CKD related complications (48). In the United States (US), the National Health and Nutrition Examination Survey (NHANES) cross sectional surveys showed prevalence of CKD stage 1-4 between 10% ((NHANES III (1988-1994)) and 13.1% ((NHANES IV (1999-2004)). The largest increase was in stage 2 (2.7% to 3.3%) and stage 3 (5.4% to 7.7%) with a superior female representation in both studies (49). A cross-sectional study conducted in Beijing (China) found that the prevalence of CKD defined by (eGFR<60ml/min/1.73m^2) and stratified by age was: 10% aged 18 to 39 years,

14.2% aged 40 to 59 years, 20.8% aged 60 to 69 years and 30.5% aged (>70 years) (50). Additionally, a cross sectional study aimed at determining the prevalence of reduced kidney function in an Australian population by the presence of proteinuria, haematuria and GFR (<60ml/min/1.73m^2), showed that in 11.2% of cases reduced GFR was detected, proteinuria in 2.4% and haematuria in 4.6% of cases (51). Furthermore, in a Norwegian study, Hallan and colleagues (52) using data from the second Health Survey of Nord-Trøndelag County (HUNT II) estimated an overall CKD prevalence of 4.7% in adults aged (≥20 years).

1.6 Natural history of CKD

CKD is a progressive disease which is independently associated with poor health outcomes. The disease remains typically asymptomatic during the earlier stages until the condition progresses to a more advanced state. Kidney function may deteriorate and get worse over months or years depending on the severity of the associated aetiologies. However, the mechanism underpinning rate of progression of CKD is largely unknown. The Framingham Heart Study Offspring Cohort (1991-1995) examined the development of CKD (eGFR <59 ml/min/1.73m^2 in women and <64 ml/min/1.73m^2 in men) after patients were given an oral glucose tolerance test and followed up for an average 7 years. The subsequent mean GFR at follow up were: normo-glycaemia (87 ml/min/1.73m^2), IGR (85 ml/min/1.73m^2), newly diagnosed diabetes (82 ml/min/1.73m^2) and known diabetes (78 ml/min/1.73m^2) (5).

The NHANES survey demonstrated a strong association between increased rates of CKD risk factors such as diabetes mellitus, cardiovascular disease, hypertension, age, obesity and severity of renal dysfunction (53). In a recent meta-analysis exploring the contribution of low eGFR and albuminuria in both a general and high risk CKD population developing ESRD and CKD outcomes, lower eGFR and higher albuminuria were both risk predictors of ESRD,

acute kidney disease and progressive CKD independent of each other and of cardiovascular disease (14).

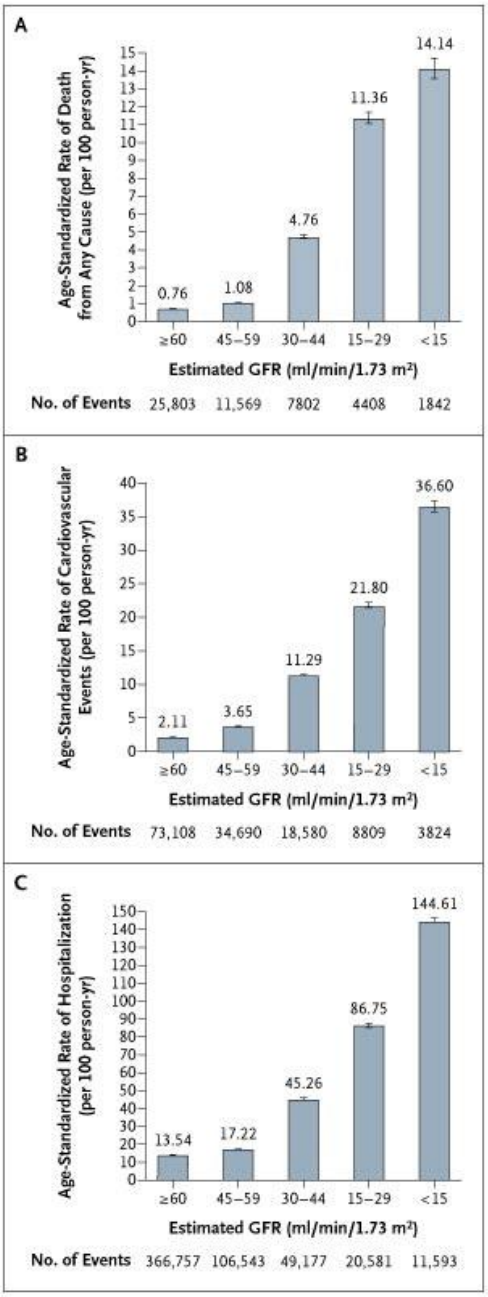
1.7 Cardiovascular risk in CKD

As CKD progresses, specific risk factors associated with renal disease and accelerated development of cardiovascular events come into play. People with CKD are at increased risk of morbidity and mortality as a consequence of cardiovascular disease. After adjusting for traditional cardiovascular risk factors such as hypertension, dyslipidaemia and diabetes, the risk of cardiovascular disease increased by approximately 2 to 4 times in individuals with impaired renal function and raised albuminuria (54).

A review by Baigent and colleagues (55) also revealed that patients with CKD stage 4-5 have a cardiovascular disease related mortality 2-4 times higher compared to the general population, whilst the risk is 100 fold higher in patients with ESRD. Studies in various populations have demonstrated that reduced eGFR and proteinuria are independently associated with an increased risk of cardiovascular disease.

In a large US community based study that used data from a renal registry, age standardised death, cardiovascular events and hospitalisation was analysed for various levels of GFR. Cardiovascular events were defined as hospitalisation for coronary heart disease, heart failure, stroke, and peripheral arterial disease. Across all studied events, a clear association between age standardised rate of events and GFR progression was apparent. As GFR declined ($\geq$60 ml/min/1.73m^2), the rate of death, cardiovascular events and hospitalisation increased in a graded pattern (Figure 4) (56).

Figure 4: Age-Standardised Rates of Death, Cardiovascular events, and Hospitalisation by decline in GFR



[1] A cardiovascular event was defined as hospitalization for coronary heart disease, heart failure, ischemic stroke, and peripheral arterial disease.

[2] Error bars represent 95 percent confidence intervals. The rate of events is listed above each bar.

[3] Reproduced with permission from (Go, A.S., Chertow, G.W., Fan, D., McCulloch, C.E., Hsu, C (2004) Chronic Kidney Disease and the Risks of Death, Cardiovascular Events and Hospitalization. The NEW ENGLAND JOURNAL OF MEDICINE, 351(13): 1296-305.

1.8 Type 2 Diabetes – A risk factor for CKD

Diabetes is a major risk factor for the development of CKD. Long term complications of diabetes often lead to either diabetic nephropathy or vascular damage which may result in kidney failure requiring dialysis or kidney transplantation. The health burden of CKD is likely to rise because the prevalence of type 2 diabetes (T2DM) has been rapidly increasing worldwide (57). A recent prospective study using data from 368 general practices (GP) contributing to the QRESEARCH database was conducted in England and Wales. A random cohort of patients aged 35 to 74 years registered with a GP practice between 1st January 2002 and 31st December 2008 and without evidence of CKD at baseline was identified. Patients had approximately 7 years of follow-up data available. The adjusted hazard of developing CKD stage (3b, 4 and 5) in people with T2DM was approximately 5 times in women (HR: 4.50; CI, 4.14 to 4.89) and 6 times in men (HR: 6.07; CI, 5.61 to 6.57) (58).

In England the estimated prevalence of CKD as a direct result of diabetes vary across population and the definition used. To establish the impact of ethnicity on the prevalence of diabetes mellitus and development of CKD, a cross-sectional study of diabetic patients (93% coded as T2DM) was carried out. The prevalence of CKD stage 3-5 was 18% in this population. The White population had a higher prevalence of stage 3 CKD than Asians (OR: 0.79; CI, 0.71 to 0.87) and Black (OR: 0.49; CI, 0.43 to 0.57). Results of this study should be interpreted with caution, as the result may not be generalisable. This study was conducted mostly in deprived areas of East London and approximately half of the population was classified as non-White (59).

The highest CKD prevalence estimate was in a study involving 17 general practices in Kent, Greater Manchester and West Surrey. Data on patient's demographics, laboratory results, diagnosis and prescriptions were extracted from routine electronic patient records. The study

revealed that 31% of patients with diabetes had CKD stage 3-5 compared to 6.9% without diabetes. Patients with diabetes were identified with a Read code suggesting a diagnosis of diabetes, the researchers however, did not categorise patients as type 1 or type 2 diabetes (60). The difference in prevalence may also be attributed to increased creatinine measurement as part of routine follow-up of diabetic patients.

1.9 Origin and background of impaired glucose regulation (IGR)

The term IGT was first introduced by the National Diabetes Data Group in 1979 (61). This was endorsed a year later by the World Health Organisation (WHO), replacing the terms "borderline diabetes" and various categories of glucose intolerance. The newly proposed definition suggests that individuals with IGR are at high risk of developing overt diabetes and are at significantly increased risk of death and morbidity due to cardiovascular disease, although it was also recognised that a proportion will revert back to normal glucose tolerance (62). IGR therefore represents an intermediate state between normal glucose homeostasis and diabetic hyperglycaemia. Individuals who are classified as IGR have a blood sugar level raised beyond normal level but it is not high enough to suggest overt diabetes (63).

1.10 Similarities and differences between IGT and Impaired Fasting Glucose (IFG)

The American Diabetes Association (ADA) Introduced IFG and/ or IGT as a disease process intermediate between normal glucose homeostasis and diabetic hyperglycaemia (64). The main difference between IFG and IGT can be explained by the marked difference in plasma glucose concentration at baseline and after a standard 75-g ingestion of oral glucose. People with IGT at baseline have similar fasting plasma glucose to those with normoglycaemia during a standard 2 hour oral glucose tolerance test (OGTT). Following glucose ingestion the plasma glucose concentration rapidly rises and fails to decline after 2 hours. On the other hand, patients with IFG have a higher fasting plasma glucose concentration at baseline

compared to those with normoglycaemia or IGT but the physiological response to oral glucose administration is normal (4).

1.11 Diagnosis of IGR

For the purpose of this review, IGR was classified as a (fasting plasma glucose (FPG) <7 mmol/l (<126 mg/dl) or a two hour Oral Glucose Tolerance Test (OGTT) of ≥7.8mmol/l and <11.1 mmol/l (140-200mg/dl), or a glycated haemoglobin (HbA1c) of 6.0 - 6.4% (42 - 47 mmol/mol) (65, 66). This condition is diagnosed using a FPG and an OGTT test. The OGTT measures a person's blood glucose level before and after a 2 hour glucose load. Prior to the test clinicians advise patients not to eat or drink for up to 8-12 hours. After the test it should be possible to determine if the person has IGR based on the amount of glucose in the blood before and after the glucose load. If a person has IGR the amount of glucose in their blood will be between (6.1 - 6.9 mmol/l) before the test and (7.8 - 11 mmol/l) two hours after the test (65, 67). The International Expert Committee (IEC) advocates the use of HbA1c assay in the diagnosis of diabetes. The report recommends an HbA1c threshold of ≥6.5% (≥48 mmol/mol) with a repeat test as confirmation of diabetes in asymptomatic individuals. Patients with a value between 6.0% - 6.4% (42 - 47 mmol/mol) HbA1c should be considered at high risk of progressing to overt diabetes (67).

The American Diabetes Association (ADA) supports the use of HbA1c as a satisfactory measurement to define patients at risk of progressing to diabetes and maintained IGR threshold as a suitable and viable measure to determine whether patients are at increased risk of developing diabetes (68). ADA further revised the cut-off value of HbA1c from 6.0% - 6.4% proposed by IEC to 5.7% - 6.4%, arguing this lower threshold has better combined sensitivity and specificity (68). In the UK, an expert advisory body convened by the Department of Health recommended that patients with an HbA1c of 6.0% to 6.4% (42 - 47

mmol/mol) should be considered at increased risk of developing diabetes and retested annually. The term high risk group was used as an equivalent to describe individuals with impaired glucose regulation which encompasses IGT and IFG (69)

1.12 UK burden of IGR

1.12.1 UK prevalence of IGR

The prevalence of IGR is poorly characterised in the UK, a number of studies were conducted in specific population with relatively small sample size and lack of standardisation and accuracy of glucose measurement to identity patients with IGR, could potentially mean that prevalence rates of IGR may not be standardised across studies (70, 71) Two London based studies in adults over the age of 40 years, reported IGR prevalences of (4.2% in men and 3.4% in women) and 4.1% in a study population of 1919 and 1040 respectively (72, 73). Furthermore, a study of 1482 adults aged 59 to 70 years conducted in Hertfordshire, reported a dramatic increase in prevalence rate (19% in men and 31% in women) (74). In contrast, a study conducted in 193 young adults aged 18 to 40 years reported a prevalence rate of 9.3% at high risk of developing diabetes and 6.2% diagnosed with IGR (75).

Impaired glucose regulation is part of the umbrella term “pre-diabetes” and is the most widely used definition of describing blood glucose level higher than normal but below the threshold for diabetes. The term prediabetes has received little support from some expert groups namely: WHO, International Diabetes Federation (IDF) and National Institute of Clinical Excellence (NICE), because of the stigma attached to the word diabetes and many people will not progress to diabetes as the term prediabetes appears to suggest. Nevertheless, many research articles still refer to impaired glucose metabolism as pre-diabetes. The prevalence of pre-diabetes in England was recently analysed from a population based cross-sectional study. Data from the years 2003, 2006, 2009 and 2011 were extracted from Health Survey for England (HSE) for adults aged 16 and above and provided a blood sample. Pre-

diabetes was defined using the HbA1c cut-off value of 5.7% - 6.4%. Individuals were excluded if they were previously diagnosed with diabetes or currently on diabetic medications. Weighted blood samples were used in the analysis. This is because HSE provides different levels of weights for analysing different variables and weighted samples allow generalisability of the adult population of England. The weighted sample size for 2003 was 7892, for 2006 it was 6385, for 2009 it was 2172 and 2011 the sample size was 3690. The proportion of individuals aged 16 and older with pre-diabetes, also known as non-diabetic hyperglycaemia increased from 11.6% in 2003 to 35.3% in 2011. In contrast, only the second most deprived quintile (quintile 4) appears to be significantly associated with pre-diabetes. The odds ratio (OR) and confidence interval (CI) for the risk of pre-diabetes in 2003 and 2011 were: 1.62 (1.26 to 2.07) and 1.45 (1.21 to 1.88) (76).

Additionally, a report published by the National Cardiovascular Intelligence Network on behalf of Public Health England (77), to determine the prevalence of non-diabetic hyperglycaemia also known as pre-diabetes or impaired glucose regulation in adults aged (≥16 years), used a combined dataset (54,644 records) from 2009 to 2013 from HSE. Non-diabetic hyperglycaemia was defined as an HbA1c value between 6.0 – 6.4% (42 - 47 mmol/mol) excluding those diagnosed with diabetes at baseline. Prevalence of non-diabetic hyperglycaemia was categorised by age group, in the age group 16 to 39 years, the prevalence of non-diabetic hyperglycaemia was 2.6%.

1.12.2 Incidence of IGR

In order to estimate the incidence of IGT/IFG in primary care, a descriptive study was carried out using the General Practice Research Database (GPRD), a large longitudinal database which provides anonymised data that can be used for research purposes. GPRD represents approximately 6% of the UK general population in terms of age, sex and geographical distribution. Patients were identified as IGT/IFG if they had a recorded measurement of

glucose in blood, urine, random blood samples or codes (READ or OXMIS) suggesting IGT/IFG. The date of the first recorded measurement or codes was taken as the index date. Patients diagnosed with diabetes or on diabetes medications before the index date were excluded. Potential IGT/IFG cases were confirmed if patients had no subsequent diabetes for at least 1 month after the index date. From a study population of approximately 2.8 million patients registered with a GP practice between 1st January 2000 and April 2005 with a minimum of 6 month follow-up, a total of 9096 patients were identified with at least one code suggesting IGT/IFG and were included in the final analysis. Participants were stratified in age band: 21.9% aged 20 to 39 years, 28.4% aged 40 to 39 years, 40.1% aged 60 to 79 years and 9.6% aged (≥80 years). The annual incidence rate of IGT/IFG increased from 17 cases per 100,000 person-years in 2000 to 31 cases per 100,000 person-years in 2004 (78).

1.13 Global burden of IGR

According to the International Diabetes Federation (IDF), an estimated 344 million or 7.9% of adults worldwide have IGR, and this number is projected to increase to an estimated 472 million or 8.4% of the population by 2030. Approximately 130 million of this population are aged 40 to 59 years and if left untreated are at higher risk of progressing to T2DM. IGR in this age group will continue to rise and is projected to reach 180 million by 2030. Furthermore, nearly one-third of all those with an IGR diagnosis in 2010 were in the age group 20 to 39 years (79).

1.14 Variation by ethnic group

The prevalence of impaired glucose regulation (IGT/IFG) was determined in a multi-ethnic cohort of 193 obese young adults aged 18 to 40 years across Leicestershire and Northamptonshire. The prevalence of impaired glucose metabolism was 32.5% in Black and minority ethnic (BME) populations compared to 14.5% in a White population (75). To

describe ethnic differences in the prevalence of IGR, a study of 1894 patients who attended a glucose tolerance test was conducted in Foleshill (Coventry, UK), reporting the following prevalence rates: Caucasians (5.7% male and 6.8% female) and Asians (9.8% male and 11.2% female) (80). Furthermore, a cross sectional study set in Leicestershire, UK, data of 3707 patients aged 40 to 74 years were analysed for the joint prevalence of diabetes, impaired glucose regulation (IGR), cardiovascular disease and CKD. Impaired glucose regulation was defined as either impaired fasting glucose or impaired glucose tolerance. Impaired glucose tolerance (IGT) was defined according to WHO criteria (2011) by a 2 hour post glucose challenge between 7.8 and 11.0 mmol/l and impaired fasting glucose (IFG) was defined as fasting blood glucose between 6.1 and 6.9 mmol/l (81). Patients were stratified by age and ethnicity. The largest cohort was from White European (76.4%) compared to South Asian (23.6%). The prevalence of IGR in White European males was 12.5 compared to 9.2% in South Asian males. Similarly, the prevalence in White European females was 10.8% compared to 9.3% in South Asian females (82).

1.15 Associations with other illnesses

1.15.1 Risk of macrovascular disease in IGR

Not many studies have comprehensively looked at cardiovascular risk factors in people with IGR. Studies have closely linked IGR with similar cardiovascular and metabolic risk profiles as T2DM (83). The risk of cardiovascular disease (CVD), ischaemic heart disease (IHD) and stroke were analysed in a large Koreans (384,795) population aged ≥20 years with IGR, followed-up for approximately 10 years. IGR was defined according to the 2003 ADA criteria of 110 - 125 mg/dL or 6.1 - 6.9 mmol/L, excluding those with coronary heart disease, cerebrovascular accident or malignancy at baseline. After adjusting for age, there was a 30% increased risk of CVD in patients with IGR (HR, 1.30; 95% CI, 1.24 to 1.35). In a recent meta-analysis, Huang and colleagues (84) examined 53 prospective cohort studies with a total

population of more than 1.6 million participants aged ≥18 years with prediabetes compared to those with normoglycaemia and risk of cardiovascular (CVD) outcomes, coronary heart disease (CHD) and stroke over a mean follow-up period of 9.5 years. Prediabetes was classified according to the World Health Organisation (WHO) guideline (FPG: 110 - 125 mg/dL or 6.1 – 6.9 mmol/L) or OGTT ≥7.8mmol/l and <11.1 mmol/l or the American Diabetes Association (ADA) lower cut-offs (FPG: 100 – 125mg/dl or 5.6 – 6.9 mmol/l) and glycated haemoglobin (HbA1c) of 5.7 - 6.4% (39 - 47 mmol/mol) according to ADA or 6.0 – 6.4% (42 – 47 mmol/mol) according to the National Institute for Health and Care Excellence (NICE) criteria. This review demonstrated a 13% increased risk of CVD outcomes (RR, 1.13; 95% CI, 1.26 to 1.30), risk of CHD and stroke increased by 10% (RR, 1.10; 95% CI, 1.18 to 1.20) and 6% (RR, 1.06; 95% CI, 1.17 to 1.20) respectively.

In The Diabetes Epidemiology: Collaborative analysis Of Diagnostic Criteria in Europe (DECODE) study group of 10 prospective European cohort studies with 15,388 men and 7126 women aged 39 to 89 years, the risk of CVD, coronary heart disease (CHD), stroke and all-cause mortality in patients with IGR and normoglycaemia where analysed, adjusted for age, study centre, systolic blood pressure, body mass index (BMI), serum cholesterol level, smoking status and sex. The aged standardised mortality in men with normal glucose tolerance (NGT) increased from 7% (CVD), 4.2% (CHD), 1.2% (Stroke), and all-cause mortality 19.3% to 8.3% (CVD), 4.9% (CHD), 1.4% (Stroke) and all-cause mortality 22.8% in men with IGR. In the smaller cohort of women with normal glucose tolerance (NGT) the mortality rate increased from 3.2% (CVD), 1.1% (CHD) and 6.9% all-cause mortality to 3.7% (CVD), 1.2% (CHD) and 9.6% all-cause mortality in women with IGR, however there was a slight decrease in mortality from stroke, the rate decreased from 1.1% (NGT) to 1.0% (IGR) (85). In a systematic review and meta-analysis of prospective observational studies performed to estimate the relative risk (RR) of developing CVD among participants with

IGT/IFG, 18 publications were identified, of which eight publications had information on the estimated RR associated with IGT. Six out of eight publications adjusted for age, smoking status, blood pressure and lipid profiles. The fixed-effects summary estimate of RR after adjustment for potential confounders was 1.20 (95% CI, 1.06 to 1.35), showing a statistically significant association (86). In addition, a meta-regression of 20 studies assessing the relationship between glucose threshold and cardiovascular risk when adjusted for risks factors such as hypertension and lipids showed a statistically significant association, however the results also showed that IGR remains an independent risk factor in the development of CVD (87).

1.15.2 Risk of microangiopathic complications associated with IGR

IGR has been shown to be independently associated with an increased risk of microangiopathic complications which include retinopathy, neuropathy and nephropathy (88).

- Retinopathy

Multiple studies on retinopathy have reported a higher risk of retinal injury/reduced visual acuity in patients with IGR. In a population based study, Nagi and colleagues (89) examined the risk of retinopathy by glycaemic status in 1219 Pima Indians aged ≥15 years. Glycaemic status was categorised as normal glucose tolerance, IGR and diabetes. Plasma glucose was measured using the World Health Organisation (WHO) criteria and retinopathy was examined by fundus photography. The prevalence of retinopathy in patients with IGR was 12% compared to 3% in those with normal glucose tolerance. Additionally, a subset of patients with IGR from The Diabetes Prevention Programme (DPP) was followed (mean follow-up: 5.6 years) for the development of diabetic retinopathy. Patients were excluded if they had a history of diabetes or were on glucose lowering treatment at baseline. IGR was defined by the following criteria [FPG between 5.3 – 6.9 mmol/l and OGTT between 7.8 –

11.0 mmol/l]. Retinopathy was defined using the Early Treatment Diabetic Retinopathy Study (ETDRS) level 20 criteria. Retinopathy was detected in 7.9% of the IGR patients (90).

- Neuropathy

Lee and colleagues (91) examined data of patients who participated in the Prospective Metabolism and Islet Cell Evaluation (PROMISE) study, an observational study aimed to investigate the association of glycaemic status and risk of peripheral neuropathy/nerve dysfunction. Glycaemic status was categorised as normal glucose tolerance, IGR and diabetes and determined according to the 1999 WHO diagnostic criteria. IGR was defined as fasting plasma glucose of 6.1–6.9 mmol/L and a 2 hour OGTT of 7.8 – 11.0 mmol/L. Data of 496 patients were assessed for the presence of neuropathy and nerve dysfunction at the 3-year follow-up examination. The prevalence of peripheral neuropathy was 49% in patients with IGR compared to 29% in those with normal glucose tolerance.

- Nephropathy

Metcalf and colleagues (92), studied urinary albumin concentration in 5467 Maori, Pacific Islander, and European workers aged ≥40 years who participated in a health screening survey of 46 companies across New Zealand. Microalbuminuria was defined as a urine albumin concentration of 29-299mg/L. IGR was determined according to WHO criteria of 2 hour plasma glucose concentrations of 7.8-11.1mmol/L. A higher prevalence of microalbuminuria was found in individuals with IGR (16.1%) compared to those with normoglycaemia (4%).

1.16 Natural history of IGR

Imbalance in glucose homeostasis usually precedes the development of diabetes mellitus. This state is usually referred to as impaired glucose regulation which encompasses IGT and IFG. Patients with IGR have been shown to be at increased risk of developing T2DM and at

increased risk of micro and macro-vascular complications independent of patients progressing to diabetes (65).

IGR is a dynamic and reversible condition. Patients with IGR may follow more than one pattern, some may revert to normal glucose tolerance while some may progress to diabetes mellitus and some may stay in the IGR range (93). The natural history of IGR has been reported in a number of studies. In a ten year prospective study of 241 individuals with IGR, 15% developed T2DM, 53% reverted to normal glucose tolerance and 22% remained glucose intolerant (94). Similarly, a 4 year prospective study conducted in 128 South African Indians with IGR at baseline, found that after 4 years, 50.4% progressed to diabetes, 24.8% persisted with IGR and 24.8% reverted to normoglycaemia (95). In a cohort study conducted in 1342 non-diabetic residents of Hoorn (Netherlands) from October 1989 to February 1992 to determine the cumulative incidence of diabetes in different categories of dysglycaemia, glucose measurements were determined according to the diagnostic criteria of WHO (1985 and 1999) and ADA (1997). The 6 year cumulative incidence of diabetes in people with IGR and NGT (normal glucose tolerance) at baseline and according to WHO (1985) diagnostic criteria was: 32.4% (111 individuals) and 3.7% (1231 individuals) respectively (96).

In a systematic review and meta-analysis of prospective studies published from 1979 through to 2004, annualised incidence rates of progression to diabetes and reversion to normoglycaemia in patients with various categories of glucose intolerance was performed. The annualised incidence of diabetes in patients with IGR was between 1.8 to 16.8%, patients with isolated impaired glucose tolerance (IIGT) varied from 4.4 to 6.4% and patients with a combination of IGT/IFG was between 10 to 15%. The meta-analysed overall risk and annualised relative risk of studies with people diagnosed with IGR reverting to

normoglycaemia after 1 year were 0.33 times as likely compared to normoglycaemic individuals (97).

Caution should be used in the interpretation of these prevalence rates, IGR has a heterogeneous pathogenesis which may be influenced by the individual or study population as well as standardisation and accuracy of the glucose measurement (71). Furthermore, overestimation of clinical diagnosis, which is almost always based on one positive result rather than repeated measurements has made it difficult to interpret the natural history of IGR in individuals or population (64).

1.17 Who is at risk of developing impaired glucose regulation?

IGR evolves gradually over time, there are several risk factors associated with a higher risk of developing IGR. These risk factors are essentially the same as risk factors similar with the development of T2DM:

- Overweight and obese

Diaz-Redondo and colleagues (98) examined data of 2,022 participants included in the Primary Health Care on the Evolution of Patient with Prediabetes (PREDAPS) prospective cohort study. The study cohort consisted of subjects aged 29 to 75 years with prediabetes at baseline compared to those without a glucose abnormality, followed for a minimum of 5 years. Prediabetes was defined by the following criteria [FPG between 100 and 125 mg/dl, and/or a glycated haemoglobin HbA1c range between 5.7% - 6.4% (39 – 47 mmol/mol)]. Data were collected by 125 general practitioners (GPs) during routine clinical consultations across Spain. Data of blood pressure, height, weight and waist circumference, hypertension, cholesterol, smoking habit and alcohol consumption were obtained during these consultations. Participants with general obesity (BMI ≥30 kg/m^2) were approximately twice as likely to have prediabetes in both men (OR: 2.51) and women (OR: 2.26) compared to

those without obesity after adjusting for age, educational level, marital status, region of residence, and family history of diabetes.

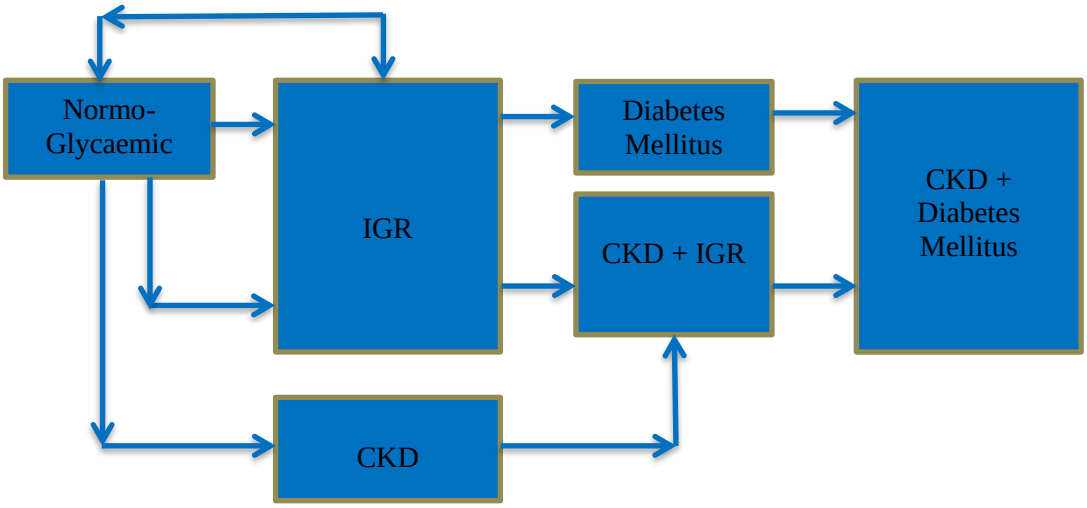
- Hypertension

Morales and colleagues (99) examined data of patients who participated in the San Antonio Heart study to determine the incidence of T2DM and IGR in hypertensive patients compared to those without hypertension. Data of 1,471 Mexican Americans and non-Hispanic Whites aged 25 to 64 years free from diabetes and IGR at baseline and follow-up for 8 years were examined. Patients with missing data, and those on antihypertensive drugs were excluded from the study. Patients with hypertension were more likely to develop IGR (25.2%) than normotensive patients (10.0%). After adjusting for age, sex, ethnicity, BMI, subscapular-to-triceps skin-fold ratio, fasting glucose and fasting insulin, risk of IGR in hypertensive patients was attenuated but remained statistically significant (OR: 1.87; 95% CI, 1.08 to 3.22).

1.18 Risk of CKD amongst young adults with IGR

In a cross sectional study among Australian adults aged (≥25 years), the prevalence of albuminuria, an early marker for the development of CKD, was 5.1% with normal glucose tolerance, 11% with IGR, 17.8% newly diagnosed and 36.2% known T2DM (100). Similarly, data from the National Health and Nutrition Examination Survey (NHANES) from 1999 through to 2006, the prevalence of CKD was found to be 39.6% in self-reported diagnosed diabetes, 41.7% in previously undiagnosed diabetes (FPG ≥126 mg/dl), 17.7% in pre-diabetes (IGR ≥100 and <126 mg/dl) and 10.6% in those without glycaemic abnormality (101). These cross sectional data are however subject to some limitations. As it is unclear whether CKD precedes impaired glucose metabolism or vice versa (Figure 5) and eGFR values were taken on only a single occasion.

Figure 5: IGR progression to T2DM/reversion to normal glucose tolerance and development of CKD



1.19 IGR trajectory and the development of CKD in young adults aged 18 to 40 years

Two epidemiological studies in Native American populations found evidence linking renal function to IGR. One found similar rates of decline in GFR over four years in persons with IGR (14% decline) and with newly diagnosed diabetes (18% decline) (102). The other found 15% of the 934 non-diabetic participants had micro-albuminuria (103). In contrast, data derived from the Framingham Heart Study offspring cohort (1991-1995) examining the development of CKD (eGFR <59 ml/min/1.73m^2 in women and <64 ml/min/1.73m^2 in men) after patients were given an oral glucose tolerance test and followed up for an average 7 years, the subsequent mean GFR at follow up were: normo-glycaemia (87 ml/min/1.73m^2), IGR (85 ml/min/1.73m^2), newly diagnosed diabetes (82 ml/min/1.73m^2) and know diabetes (78 ml/min/1.73m^2) (5).

1.20 Use of electronic health/patient record databases for research

Large health care databases are increasingly being used to facilitate medical research in the UK and around the world. Recent UK government policy proposes that by 2020 patient records should be virtually paperless (104). Countries such as Canada, the United States, Netherlands, Spain and some Scandinavian countries have established electronic health

record repositories which are increasingly being used for medical research (105-109). The UK took an early lead in developing high quality patient records repositories, particularly primary care databases which have been assessed both for their validity and quality through numerous studies (110). To further harness the capability of a diversity of existing databases the UK government announced their support to the creation of large data warehouses (also known as big data) which is part of the government 'eight great technologies' strategy (111). Research institutes such as the Farr Institute of Health Informatics Research (112) and the Alan Turing Institute (113) received significant investments to support the big data strategy and drive high quality research linking routinely collected electronic health data to other national databases (114).

There are several large and many smaller primary care databases currently in use in the UK, the three most prominent systems widely used for epidemiological research are: the Health Improvement Network database (THIN) (115), Clinical Practice Research Datalink (CPRD) (116) and QResearch (117). These databases flourished because of the almost complete adoption of computerised patient records in general practices across the UK since the 1990's. This was further enhanced by the Government incentivisation scheme the Quality and Outcomes Framework (QOF) (118), requiring General Practitioners (GPs) to record more information in electronic patients records and therefore improved the quality of record keeping. These databases are powerful tools for researchers if used appropriately.

1.21 Summary

It is widely accepted that the risk of CKD is elevated in patients with diabetes mellitus. It is far less certain whether this risk is also present in patients with IGR. As many studies have used a single determination of glycaemic status at baseline, it is not clear whether the risk of developing a CKD event is confined to people with IGR who progress to overt diabetes or whether the risk is still increased among people with IGR even if they never develop

diabetes. Also many studies do not particularly include young adults aged 18 to 40 years in their analyses.

Due to these limitations it is inappropriate to extrapolate rates and relative risks to this narrower age group of individuals and in particular to primary care where the majority of decisions for early prevention are made. It is therefore important to have a reliable estimate of the incidence of CKD in this cohort of individuals.

1.22 Research aims and objectives

- Aim:

The overall aim of this study is to elucidate whether the presence of IGR is associated with an increased risk of CKD by comparing the risk of CKD in individuals with IGR to those without IGR.

- Objectives:
    - Determine the incidence and period prevalence of IGR in young adults aged 18 to 40 years
    - Determine the incidence of CKD in young adults aged 18 to 40 years with and without IGR and after adjustment for confounders
    - Determine the incidence of CKD by category (stages 1 – 2 /3 – 5) in young adults aged 18 to 40 years with IGR
    - Determine CKD predictors in IGR to identify likely risk factors and determine their value in determining incident CKD in young adults aged 18 to 40 years with diagnosed IGR

CHAPTER TWO
PRIMARY CARE DATA FROM THE HEALTH IMPROVEMENT NETWORK (THIN) DATABASE

2.1 Background

In the United Kingdom (UK) almost the entire population is registered with a general practitioner (119), with some individuals registered on a practice list contributing several decades of data which are captured in computerised medical records. A large volume of electronic patient records have been collected since the 1990's (120). Some practices contribute anonymised electronic patient records to large research databases. The Health Improvement Network (THIN) database is one of several large UK databases arising from general practice electronic patient records which are being used for health services research (121).

Other large databases which provide patient data from across the UK include Clinical Practice Research Datalink (CPRD) (116), QRESEARCH (117) and ResearchOne (122) databases. CPRD (formerly known as General Practice Research Database (GPRD)) is managed by the Medicines and Healthcare products Regulatory Agency (MHRA) and as of July 2013 contained anonymous data of over 13 million patients (4.4 million active) from 674 practices. The database represents approximately 6.9% of the UK population (116). QRESEARCH database on the other hand contained 13 million pseudonymised patient records from 754 practices in May 2014, representing approximately 7% of the UK population (117). Additionally, ResearchOne holds data of 6.3 million patients with 4.9 million active from approximately 413 practices (122). Furthermore, in 2014 THIN database contained approximately 12 million patients record from over 600 practices contributing to the database, covering approximately 6% of the UK population (123). Participating practices

voluntarily contribute data to these databases. Patient records are uploaded to research databases on a regular basis. THIN data for example are collected on a monthly basis. For practices using Vision (THIN and CPRD) or EMIS (QRESEARCH) software, the electronic records are pseudonymised, in that the name, address, date of birth, post code, NHS number and non-structured data such as scanned hospital letters are removed from patient records before they are collected. Some free text is de-personalised by the data custodian (e.g. anonymisation of free texts by THIN) and re-inserted back into the research database before it is made available for research purposes. Records of patients are then aggregated into databases containing millions of patient records from across the UK. Patient records are organised in individual files by a clinical software company providing a longitudinal track record for each patient allowing for studies examining incidence of disease or of prediction model design. Access to these records is subject to approval from the respective scientific review board. Patients are given the option to have their de-personalised medical records collected for scientific research purposes, therefore patients consent are not required when working with anonymised patients records in the THIN database.

Recent addition to primary care research databases involves linkage to other databases such as the Hospital Episode Statistics (HES) database, the National Cancer Data repository and Office for National Statistics (ONS) mortality data (117, 124, 125). For a subset of practices contributing to the THIN database, patient level data can now be linked to the HES database (126).

2.2 The Health Improvement Network (THIN) Database

The Health Improvement Network (THIN) is a large primary care database containing longitudinal anonymised medical records of patients. In September 2014 the database contained data from 611 practices across England, Wales, Scotland and Northern Ireland. As

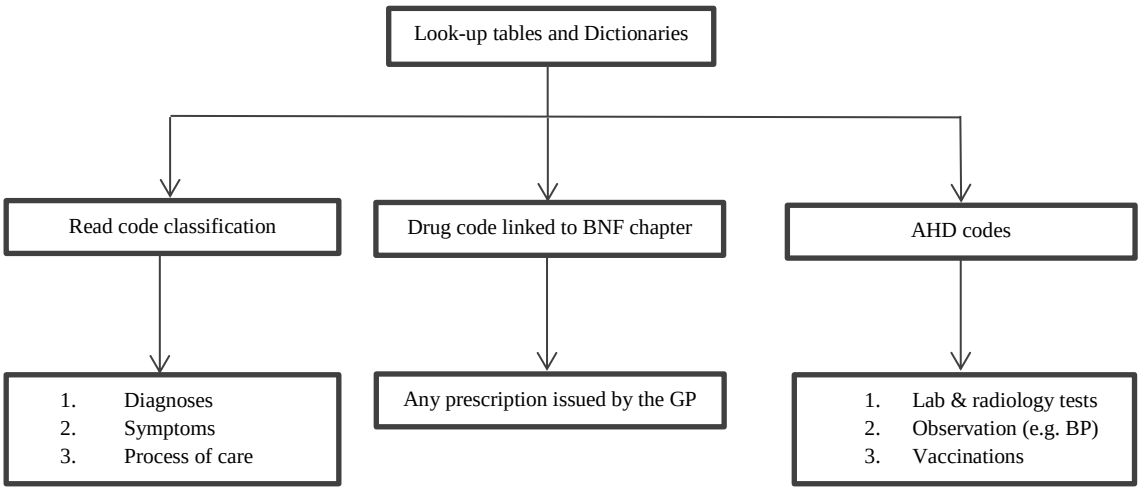
of September 2014, the database holds a total of 13 million patients of whom 3.6 million are actively registered with a general practice equivalent to more than 85 million person years of computerised data covering approximately 6% of the UK population (123). THIN database is jointly managed by IMS Health Real World Evidence Solutions (http://www.epic-uk.org/index.html) and In Practice System (InPs). THIN was developed in 2002 by the Epidemiology and Pharmacology Information Care (EPIC), a non-profit company. Three years later CSD Medical Research UK acquired EPIC, which in 2015 became part of IMS Health Real World Evidence Solutions (127).

Vision software is an electronic medical records computer system used to record the details of each GP consultation and any other information entered by the practice outside face-to-face consultations (e.g. lab results, hospital letters) (128). All general practices affiliated with the THIN database use Vision software.

Upon joining THIN, a practice undergoes a full data collection procedure allowing current and retrospective data to be collected and sent to IMS Health, the data provider. Following this, subsequent data are automatically downloaded from the practice on a monthly basis ensuring minimal disruption to the practice (123, 129). From this data, THIN generates four standardised (patient, medical, therapy and additional health data) and one linked (postcode related variable) file per practice. These data are then passed on largely unaltered to provide fully coded records. Some additional information is added in THIN at this point, for example on the likely validity of the patient registration and de-registration dates. This is important when calculating follow-up period (123). These data are processed to provide fully coded records of patient demographics, laboratory data, prescription, medical diagnoses, additional health data (e.g. smoking status, physical examination), ethnicity, socio-economic status (based on postcode of residence) and environmental indices (air quality) (Table 3) (130). The

coded data are interpreted using look-up tables and dictionaries. Diagnoses are coded using the version 2 (5-byte) Read-codes (131) hierarchical classification, which include codes related to diagnoses, laboratory and radiology tests, observation and process of care (e.g. referral). Drug prescriptions issued by the GP are recorded in the database by Gemscript coding system, managed by RESIP UK, a sister company of InPs (132). Drugs are stored using their generic names and can be linked to the British National Formulary (BNF) (Figure 6) (133).

Figure 6: Interpretation of coded data



THIN links laboratory tests and clinical measurements electronically to a coding system. Vision software allows additional information to be recorded by the practice, alongside the coded data. This is known as free-text, and may contain additional information such as additional observations, symptoms, plans for further investigations or information from scanned hospital letters. These data are uploaded from practices but are only provided to researchers after anonymisation. Currently over 30% of free texts records have been anonymised, but additional individual records can be anonymised on request at additional cost to the researcher (123).

Table 3: Information in the Health Improvement Network (THIN) database

Files	Information contained
Patient	Demographics including: • Date patient registered at practice • Date patient left practice • Death date • Year of birth • Gender • Household identifier number **THIN data does not include the following**: name; exact address or postcode; exact date of birth; NHS number
Medical	Diagnosis: • All conditions and symptoms coded by the GP/Nurse • Information on referrals to secondary care, including the speciality of the secondary care service • The GP/nurse may summarise details of prescriptions from ongoing outpatient specialist care or over-the-counter drugs • Secondary care information and other related information received by the practice can be entered including: ❖ Details on hospital admissions ❖ Discharge medication and diagnosis ❖ Outpatient consultation diagnosis ❖ Investigation and treatment outcomes ❖ Ethnicity of patient
Therapy	Prescription: • GP and in some cases nurses may issue prescriptions to patients using Vision, all prescribing is logged into the system automatically. • Acute treatments and medicines for a chronic condition can be temporally linked with a symptom or diagnosis although this is not always used by practices
Additional Health Data (AHD)	• THIN Data contains information on lifestyle and health factors such as smoking and alcohol intake. • Tests and laboratory results • Electronically linked to pathology laboratories and test results received electronically • Immunisation
Socio-economic data	The majority of patients are linked to postcode-based socioeconomic, ethnicity and environmental indicators. The data are based on the patient's postcode. Individual socioeconomic data are available in the form of the Townsend quintile
Free texts comments	Additional anonymised practice commentary, linked to records in the Medical file, may be available as free texts

Source: http://csdmruk.cegedim.com/our-data/data-content.shtml [accessed 21/01/2016]

2.2.1 Acceptable Mortality Recording (AMR)

AMR denotes the earliest period when mortality data were optimally recorded by practices which was consistent with ONS (134). Prior to changing to Vision software many practices were recording mortality rate using the Value Added Information Medical Product (Vamp) software. Immediately after changing to Vision mortality recording was inaccurate (immortal

periods) and raises a few data quality issues as practice mortality rate is determined from mortality data and it also affects the denominator, which affects all calculated incidence and prevalence rates. Patients who have died cannot get diagnosed with diseases. Discrepancies occurred because practices may have excluded patients during the changeover if patients had died before they were transferred to the new system. Therefore, the mortality rate was reduced when transferred and gradually increased once mortality rate was recorded in Vision. Furthermore, some practices split and only live patients were transferred to the new practice. In these cases data conversion occurred only on live patients (123). One study examined mortality recording in the THIN practices compared to a similar UK national death rate statistics. A marker date was derived, the Acceptable Mortality Recording (AMR) denoting the date when practices were recording mortality rate consistently and at a similar rate to the UK national statistics according to the age and gender structure of the practices (134). This is a marker for data quality in the THIN database. For this thesis it was mandatory to have an AMR prior to the index date (patients without CKD at diagnosis of IGR to ensure that practices were not under reporting important outcome and not including patients who have died). It is important to a have consistent record of AMR because mortality recording directly affects follow-up time (person-time of follow-up before developing CKD) and minimises the risk of immortal period (when death on further diagnoses occurred but were not recorded)

2.2.2 Electronic transfer of laboratory results

Laboratory test results are generally coded and stored in the patient clinical data record. Over 75% of practices contributing to THIN are electronically linked to pathology laboratory allowing all tests results to be automatically coded and entered onto the system (130). Some practices still use paper based recording, these practices manually code their results. Electronic transfer and recording of tests results would have a direct impact on the completeness of data for example, estimated glomerular filtration rate (eGFR) recording (an

important marker of CKD) would have improved since the introduction of electronic links with pathology laboratory (123).

2.2.3 Vision installation date

Prior to implementing the Vision management software many general practices were using the Value Added Information Medical Products (VAMP) clinical system to record clinical events. Data recorded using VAMP were reasonably complete regarding clinical illnesses and prescribing (135). The Vision-date indicates when a practice converted to InPs Vision software for the first time. This is the date from which patients records are likely to be complete. When gaps in data recording are identified it is usually problematic to ascertain whether there was a data issue (123). Therefore, it is advisable to use the Vision-date as one parameter in deciding the practice start date which is important when follow-up time is measured, for example calculating incidence rate of CKD for this study.

2.3 Data integrity and ethical approval

THIN database obtained ethical approval for data collection from the South East Multicentre Research Ethics Committee (reference number: 07/H1102/103). Research undertaken using the database is therefore approved provided it has undergone independent scientific review. Researchers requiring access to THIN have to submit a study protocol justifying access and gain approval from one of THIN's Scientific Advisory Committees. Such committees review study protocols and provide input whether it is an amendment or rejection of the study. The protocol for this study was reviewed and approved (study reference: 14-038) by a THIN Scientific Review Committee (SCR) in 2014.

THIN runs a set of validation checks (e.g. record of year of birth, registration status and death date) to ensure data integrity for each patient. A flagging system is used to indicate whether the data has passed these validation checks and patients have an acceptable record. For

example a transfer out date is provided in THIN as a separate field, indicating whether a patient has left the practice or died. A field providing information on death has also been created in THIN, this field allows researchers to search for information available on death from patients anonymous death certificates or from anonymised free texts comments (123).

2.4 Data quality and limitations

2.4.1 THIN representativeness

Data in THIN has been shown to be broadly generalisable of the UK population in relation to demographics and disease distribution. A study by Blak and colleagues (136) assessed the representativeness of the THIN data in terms of demographics, prevalence of some of the major chronic conditions, deprivation and death rates with UK national statistics and the 2006/2007 Quality and Outcomes Framework (QOF) prevalence data. The age and sex distribution in the THIN population were similar to the overall UK population, although there are marginally fewer people aged (<25 years) identified in THIN and the male population matches less well than the female population. Prevalence of the major QOF conditions (e.g. CKD, Diabetes and Hypertension) was shown to be similar to the national estimates. For example THIN crude prevalence of CKD was 2.5% compared to 2.3% nationally. The QOF prevalence in THIN for some of the study variables compared to the QOF national estimates are given in Table 4 (136).

Table 4: Prevalence of QOF related chronic conditions in THIN compared with UK QOF data for 2006/2007

Chronic conditions	THIN QOF %	UK QOF %
CKD	2.5 (2.5-2.5)	2.3
Diabetes	3.5 (3.5-3.5)	3.7
Hypertension	12.7 (12.6-12.7)	12.6
Atrial Fibrillation	1.4 (1.4-1.4)	1.3
Heart Failure	0.9 (0.9-0.9)	0.8
Stroke/TIA	1.9 (1.9-1.9)	1.7

TIA, transient ischaemic attack; 95% confidence interval given in parentheses

Adapted from: Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Inform Prim Care 2011;19(4):251-5.

Patients captured in the THIN database appear to be living in more affluent areas (23.5%) compared to the national average (20%). In 2006 death rates in THIN were compared to the national rate, the age-gender and deprivation adjusted death rate in THIN was 9.08 per 1000 population compared to 9.4 per 1000 population nationally. It should be noted that adjustment for deprivation when estimating death rates in THIN resulted in closer estimates to the national death rates (136).

2.4.2 THIN validity in predicting CKD

General practices contributing to the THIN database receive regular validation reports quantifying the completeness and accuracy of data recording. These feedback reports from THIN are designed to improve the quality of recording and reduce data omission (123). In a cross-sectional study Denburg and colleagues (137) examined recording of biochemistry results at general practice level using the THIN database. The aim was to accurately identify patients with moderate to advanced CKD determined by stage 3-5 using serum creatinine record and a list of 45 Read codes (indicative of moderate to advanced CKD) in the absence of complete laboratory results. Renal function stage 3-5 was measured by an eGFR<60ml/min/1.72m^2 on at least two occasions separated by a period of more than 3

months. The study demonstrated that THIN database is a reliable resource providing accurate data that can be used to identify patients with moderate to advanced CKD in large population. The list of 45 Read codes was shown to accurately identify CKD with excellent specificity (98.2% correctly Read coded without the disease) and high positive predictive value (PPV) (88.9% with a Read code indicating CKD 3-5 have the disease) but with low sensitivity (48.8% correctly Read coded diagnosis of CKD 3-5).

2.4.3 Accuracy and completeness of primary care data in the UK

The validity of routinely collected primary care data has been demonstrated in three systematic reviews (138-140). These reviews investigated a range of methods used to validate diagnoses recorded in one of the major UK primary care databases. The validation methods used a range of approaches including: questionnaires to patient's general practitioners, additional data from the general practice medical records and comparing prevalence rates from external sources such as the (Morbidity Statistics from General Practice (MSGP4) or the Doctors Independent Network (DIN) databases. With respect to some of the variables of interest (Diabetes, cardiovascular disease) to this thesis, one review (138) analysed papers reporting validation of clinical diagnoses derived from the General Practice Research Database (GPRD) found that these were well recorded (positive predictive value PPV > 90%), however atrial fibrillation another variable of interest is not so well recorded (PPV <80%). None of the reviews reported evidence of validity of CKD diagnoses, the outcome of interest of this thesis.

Herrett and colleagues (139) reviewed 212 publications to identify methods used to validate diagnoses in the General Practice Research Database (GPRD). A total of 357 validation methods accounting for 183 diagnoses were identified. These validation methods were classified as either internal (diagnostic codes, review of anonymised free texts and sensitivity analysis) or external (Questionnaire to GP, request to GP to provide anonymised medical

records and comparison of incidence/prevalence with external UK - based data source). Study findings were extracted from individual paper and categorised into disease groups. The median proportion of cases identified in the disease group of interest were: 88% of cases with endocrine, nutritional and metabolic syndrome (e.g. impaired glucose tolerance/impaired fasting glucose or diabetes) and 85% of cases with a circulatory system disorder (e.g. stroke). It should be noted that although a large number of cases were confirmed most studies in this review only considered positive predict values (PPVs).

Anandarajah and colleagues (141) collected routinely collected data from 12 UK practices and found that nearly 5% of the registered population had a record of reduced kidney function determined by an estimated glomerular filtration rate (eGFR) of <60ml/min/1.73m^2 corresponding to CKD stage 3-5. However, only 3.6% of this population had a Read code allowing identification of CKD. This would suggest that a Read code of CKD alone may not be sufficient to adequately identify patients with CKD in general practice database. A Read code of an estimated glomerular filtration rate (eGFR) or record of a measurement should also be included in addition to a Read code of CKD.

2.4.4 Quality and Outcomes Framework (QOF) impact on the recording of clinical information in primary care

The Quality and Outcomes Framework (QOF) was introduced in 2004 as part of the General Medical Services (GMS) contract. The objective of QOF is to provide incentive and reward practices for the provision of quality care and reduce variation in performance across practices (118). A set of achievement measures known as indicators were developed to ascertain practices level of achievement. These achievement indicators were initially organised into four domains (clinical, organisational, patient experience and additional services). In 2010 the QOF business rules was taken over by the Health and Social Care

Information Centre (HSCIC) from the NHS Employers and the NHS Connecting for Health (142).

In 2013 practices participating in the QOF scheme were required to document patients with chronic conditions annually. Practices were required to maintain a register of indicators for which QOF points were awarded and payment calculated. For example points were allocated for maintaining a register of adults with CKD stage 3-5, patients receiving treatment for proteinuria and annual testing of patients for protein-creatinine-ratio (PCR) (142). This register contains details of most patients with CKD in the UK (143). These are important clinical indicators with respect to the diagnoses of interest for this study. The introduction of QOF may have driven practices to improve coding of clinical information in primary care databases.

2.5 Advantages and limitations of THIN to this study

2.5.1 Strengths associated with the study

One of the key advantages of the THIN database is its size. The database holds over 3 million active patient records from over 600 practices throughout the UK. The database includes men and women of all ages, with varying health status, including conditions such as CKD. This allows study findings to be widely applicable and generalisable to the UK population. Additionally, estimation of incidence rates of CKD is permitted because of the large number of patients contributing to the THIN database, which may be otherwise impossible to study because of the potentially lower incidence of CKD in the 18 to 40 age group (123). Furthermore, as data on exposure (IGR) were collected prospectively THIN database is less likely to be subject to recall bias.

The THIN database contains longitudinal data for some patients dating as far back as 1985, making it possible to have long term patient follow-up, allowing person time to be calculated

and occurrence of new cases documented. Another advantage of THIN is the non-interventional way data are collected, allowing researchers to have a wide scope in terms of study design. Furthermore, the database is continually updated allowing immediate selection of both cases and controls from the same source population. THIN also permits identification of exposures and outcomes in various ways, for example, in addition to a Read code, laboratory results are also available to confirm diagnosis (123). A number of studies have validated the THIN data, including records on diagnoses, symptoms and prescriptions (144, 145).

2.5.2 Limitations

THIN database is susceptible to some limitations. Computerised data collected for the THIN database are collected during routine clinical consultation and if patients do not consult, then no data can be collected. Although information collected by General Practitioners (GPs) is expected to be complete some information may be less complete in terms of important study variables compared to a study designed specifically for research. These data are collected for clinical purposes and not for research and GPs do not collect all information on all patients (123). Another limitation of the THIN database is the lack of data available for some patients on important confounders such as smoking, height and weight particularly in the early years of practices recording data. Additionally, recording of information is biased because GPs selectively record information in patients where they think it is clinically relevant, for e.g. weight is more often recorded if the patient is overweight (123). These are important variables in relation to this study. However, since the introduction of QOF in 2004, recording of these have been improving.

Furthermore, before computerisation of practice records and laboratory linking, historic laboratory results were entered manually by practice personnel and in some cases only abnormal results were entered for inclusion into the THIN database (123). Since patients and

practices can enter and exit THIN database at any time, some patients may have short follow-up periods, reducing the likelihood of these patients developing the outcome of interest (CKD). To mitigate this limitation, this study opted for an open cohort study design, allowing patients lost to follow up to be replaced over the study period (2000 - 2015). Some patients may join another THIN practice and contribute historical data twice. However, for those patients follow-up will cease at their original practice and a transferred out date allocated. Patients will be allocated new registration dates and registration status of the new practice (123).

CHAPTER THREE
A SYSTEMATIC REVIEW EXPLORING THE EFFECTS OF IMPAIRED GLUCOSE REGULATION (IGR) ON INCIDENCE OF CHRONIC KIDNEY DISEASE (CKD) IN YOUNG ADULTS

3.1 Introduction to chapter

This chapter reports a systematic review of cohort and case-control studies to elucidate whether the presence of impaired glucose regulation (IGR) is associated with an increased risk of chronic kidney disease (CKD) by comparing the risk of CKD in individuals with IGR to those without IGR.

3.2 Background

Chronic kidney disease is a long term condition attributed to the kidneys inability to effectively remove waste product and excess water from the body. CKD is characterised by the presence of kidney damage and/or a gradual loss of kidney function (eGFR) over time (146). Diabetes is a major risk factor for the development of CKD. Long term complications of diabetes often lead to either diabetic nephropathy or vascular damage which may result in kidney failure requiring dialysis or kidney transplantation. A prospective cohort study conducted in England and Wales found the hazard of developing CKD in patients aged 35 to 74 years was 5 times higher in women and 6 times higher in men with diabetes than non-diabetics (58).

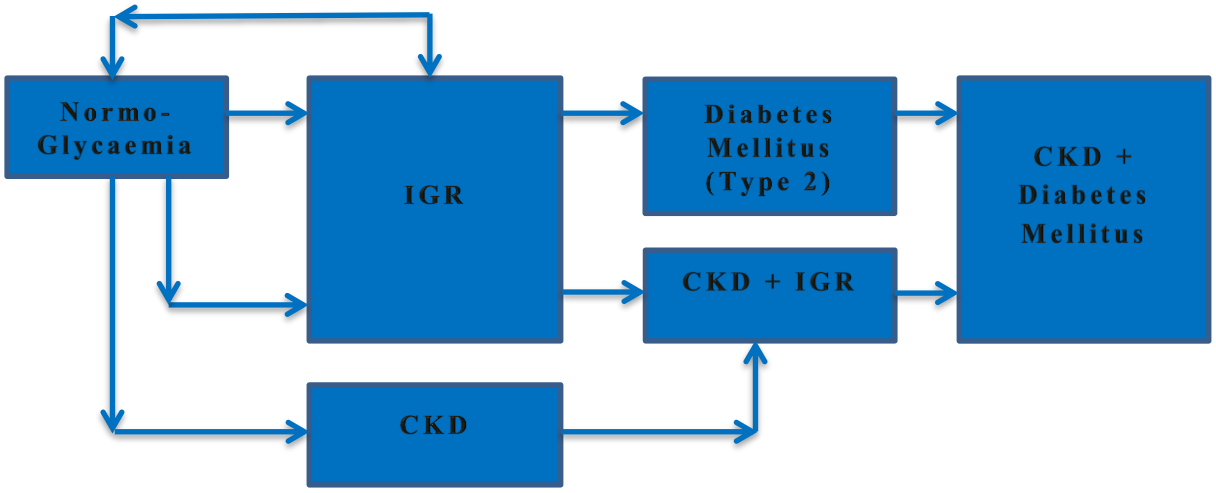
Pre-diabetes indicates both IFG and IGT collectively known as (impaired glucose regulation [IGR]). Individuals with IGR have a blood glucose raised beyond normal level but not high enough to suggest a diagnosis of diabetes (63). It is not clear whether risk of developing CKD

is elevated in patients with IGR or whether any increased risk only occurs after patients develop diabetes. Furthermore, the risk of young adults aged 18 to 40 years with IGR developing CKD is not well characterised. There is some evidence that the incidence of CKD is elevated in individuals with IGR, but this is confined to specific populations. Watanabe and colleagues (7) examined the association between metabolic syndrome and the incidence of CKD in a large prospective cohort study of Japanese adults aged ≥20 years without CKD, antihypertensive drug, diabetes or CVD at baseline. The incidence of CKD adjusted for age and sex was approximately twice higher [HR: 1.94 (95% CI, 1.06 to 3.54)] in individuals with IGR compared to those with normoglycaemia. Cross-sectional studies show that albuminuria an early marker of CKD was approximately 3 times more common in IGR than those with normoglycaemia (100). Furthermore, cross-sectional data shows some association between increasing blood glucose and decline in kidney function (147-149). Cross sectional studies are however subject to some limitations, as it is unclear whether CKD precedes impaired glucose metabolism or vice versa.

3.3 Study aims and hypotheses

To conduct a systematic review and if possible a meta-analysis of cohort and case-control studies to examine whether the presence of IGR in young adults aged 18 to 40 years is associated with an increased risk of CKD by comparing the risk of CKD in individuals with IGR to those without IGR, and also to evaluate whether any increased risk occurs only after they develop T2DM (Figure 7).

Figure 7: Progression of IGR to T2DM or reversal to normoglycaemia and development of chronic kidney disease



3.4 Methods

Established guidelines for reviews were used to inform the search strategy, selection of studies, assessment of risks of bias and reporting of results (150, 151). The review protocol was published (152) (Appendix 1). The full review has been peer-reviewed and published (Appendix 2).

3.4.1 Eligibility criteria

- Types of participants and comparison group

This review includes studies where some participants are aged 18 to 40 years and results reported separately in this age group without a diagnosis of type 1 and type 2 diabetes but with IGR, “Pre-diabetes” or “Pre-diabetic state”. IGT/IFG can be referred to as pre-diabetes; (153) or metabolic syndrome where IGR is part of the metabolic syndrome. The comparison group was either participant with normoglycaemia or diabetic participants. For the purpose of this review, IGR was classified as a (FPG <7 mmol/l (<126 mg/dl) or an OGTT ≥7.8mmol/l and <11.1 mmol/l (140-200mg/dl), or glycated haemoglobin (HbA1c) of 5.7 - 6.4% (39-47 mmol/mol) and IFG was defined as FPG of 5.6 – 6.9 mmol/l (100 – 125mg/dl) (154)

- Participants and outcomes – cohort studies

This review includes any cohort studies where some participants are aged 18 to 40 years and results reported separately in this age group with (1) IGT/IFG (exposed group) but without a diagnosis of type 1 or type 2 diabetes compared to participants without glycaemic abnormality (comparator) (2) IGT/IFG but without a diagnosis of type 1 diabetes compared to participants with T2DM. Participants were free from CKD at baseline. A broad range of measures were used to ascertain CKD (outcome). This included eGFR stages 3A, 3B, 4 and 5; albuminuria; albumin creatinine ratio (ACR ≥2.5mg/mmol or ≥30mg/g), protein creatinine ratio (PCR ≥45mg/mmol or ≥300mg/g), serum creatinine (SCr 1.0mg/dL or ≥50 μmol/L), proteinuria (≥1+) and creatinine clearance (CrCl ≥60ml/min). Studies reporting mean changes in continuous variables (e.g. eGFR) were included and findings summarised separately. Studies reporting a single measure instead of two measures of eGFR or only by any of the above measures were included. Measures of association (HR, OR, IRR and RR) and their 95% confidence interval were extracted and reported or calculated using data derived from the publications.

- Participants and outcomes – (case – control) studies

This review also includes any case-control studies where some cases are aged 18 to 40 years with an incident diagnosis of CKD (the outcome of interest) by any of the above definitions and controls without a diagnosis of CKD and results are reported separately in this age group. The frequency of previous IGT/IFG (exposure to IGT/IFG) was compared to either the frequency of normoglycaemia (unexposed) or to the frequency of diabetes (an alternative exposure). There was no restriction on the length of participant follow-up.

3.4.2 Information sources and searches

The following electronic databases were systematically searched with no language restriction from inception to January 2015: MEDLINE, Cumulative Index to Nursing and Allied Health

Literature (CINAHL), EMBASE, PubMed, Database of Abstracts of Reviews of Effects (DARE), Cochrane Database of Systematic Reviews (CDSR) and Trip Database. Furthermore, ongoing studies, scientific literature and abstract proceedings were identified by searching the following databases: ClinicalTrials.gov, Cochrane Renal Group specialised register, Renal Registry Database, British Renal Society, Renal Association, American Society of Nephrology, World Congress of Nephrology, Diabetes UK Conference, Primary Care Diabetes Society Conference and Zetoc. A comprehensive search of the Conference Proceedings Citation Index (CPCI) was also carried out. Search of these databases spanned from January 2011 to January 2014 as it is likely that studies would have been completed and published. Grey literature databases, such as Grey Literature Report, OpenGrey, PubliCat and ScienceDaily.com were examined. Google Scholar was also explored; a scoping search revealed that the most pertinent articles were found in the first ten pages of the searches. Open access theses and dissertations were retrieved from the ProQuest Dissertation Thesis Database and thesis.com. The Science Citation Index (SCI) were used to scan and track study titles. The search strategy is shown in Appendix 3. A sensitivity search of the above databases was also carried out excluding diabetes as a required search term from the initial search strategy to ensure that all the relevant studies comparing IGR (exposed group) to participants without glycaemic abnormality (comparator) are identified (Appendix 4).

3.4.3 Study selection

Two reviewers independently reviewed all titles and abstract of the search results in two phases. First the retrieved titles and abstracts were reviewed to identify relevant studies. Then the full texts of retrieved studies were read to determine eligibility. Any discrepancies or difference in opinion were resolved by consensus or by involving a third reviewer. An inclusion criteria checklist (Appendix 5) was developed based on studies eligibility criteria piloted on 5 papers.

3.4.4 Data collection process

A data extraction form has been designed based on Hayden and colleagues framework (155); this is described in more detail in Appendix 6. This form was iteratively developed and piloted on known papers. The form has been designed to focus on population, comparator, outcome and study design. Data extraction was conducted by one reviewer and checked by another for studies identified through the screening phase. Errors in data extraction were discussed and amended as appropriate. For missing data, authors were contacted for clarification. A Microsoft Excel sheet was used to manage data extraction.

3.4.5 Quality assessment

Study quality was assessed according to a modified tool based on the Ottawa-Newcastle scale (NOS) (156). Risk of bias was assessed on the following domain: (1) sampling, (2) outcome measurement, (3) attrition, (4) analytical method and (5) confounders (Appendix 7). A composite score was not provided, instead a risk of bias of 'yes' indicating adequate data were provided, 'no' if data were provided but did not meet the criteria for that domain and 'unclear' potentially at high risk of bias (157).

3.4.6 Publication bias

If eligible studies are identified, the Begg's(158) and Egger's (159) regression tests will be carried out to detect publication bias. At least 10 studies will be needed to sufficiently detect publication bias (160). Studies will be grouped according to effects measures and reporting risk of CKD determined by any of the following measures (Estimated Glomerular Filtration Rate (eGFR), albumin creatinine ratio (ACR), proteinuria ≥1, serum creatinine (SCr) and creatinine clearance (CrCl) levels).

3.5 Data synthesis

If eligible studies are identified for future update of this systematic review the following data analysis criteria may apply. Studies will be grouped according to similar effect measures and CKD markers. Adjusted and unadjusted risk estimates of CKD will be combined and pooled estimates determined. A random effect model will be used to ensure studies providing different effect sizes are represented in the summary estimate. Heterogeneity will be assessed using the Chi-squared ($\chi 2$) test with a p-value (<0.05) considered statistically significant. Chi-squared ($\chi 2$) test is poor in detecting heterogeneity, especially with small number of studies (161). Therefore the (I^2) statistics with 95% confidence interval (CI) will be used to assess the percentage variance between studies. The magnitude of heterogeneity will also be quantified using τ^2 (tau-squared) to estimate the in between study variance. Studies where effect measures cannot be combined, results of these studies will be reported narratively. Furthermore, studies reporting changes in eGFR during progression of renal disease, the standardised mean difference will be reported among subjects with IGT/IFG compared to normoglycaemia.

3.6 Results

3.6.1 Search results

The initial database searches identified a total of 5568 studies. After duplicates were removed 5478 records remained. After scanning titles, 90 citations potentially met the inclusion criteria. These were reviewed in detail (full text) and 19 cohort and no case-control studies were selected for further review for potential inclusion. A summary of the overall quality of the 19 studies which were assessed for potential inclusion and the confounding factors adjusted in the multivariate analyses are provided in Table 5 and

Table 6. A PRISMA study flow diagram of included and excluded studies is provided along with reasons for exclusion (Figure 8). The sensitivity database searches identified a total of

14,585 studies. After duplicates were removed 13,982 studies remained. After scanning titles, 114 potentially relevant articles were identified. These were reviewed in detail and 26 articles were identified, nineteen were the originally identified articles and 7 additional articles were identified, six cohort and 1 case control. A summary of the overall quality of the 7 studies which were assessed for potential inclusion and the confounding factors adjusted in the multivariate analyses are provided in Table 5. Review of the 7 studies identified that only one study reported separate data for persons with IGR aged 18 to 40 years and risk of CKD compared to T2DM. The remaining 6 studies reported risk of CKD in IGR compared to normoglycaemia or T2DM at all ages. A study flow diagram of the included and excluded studies is presented in Appendix 8 following the sensitivity search excluding diabetes from the initial search strategy (Appendix 3).

Figure 8: Study selection and reasons

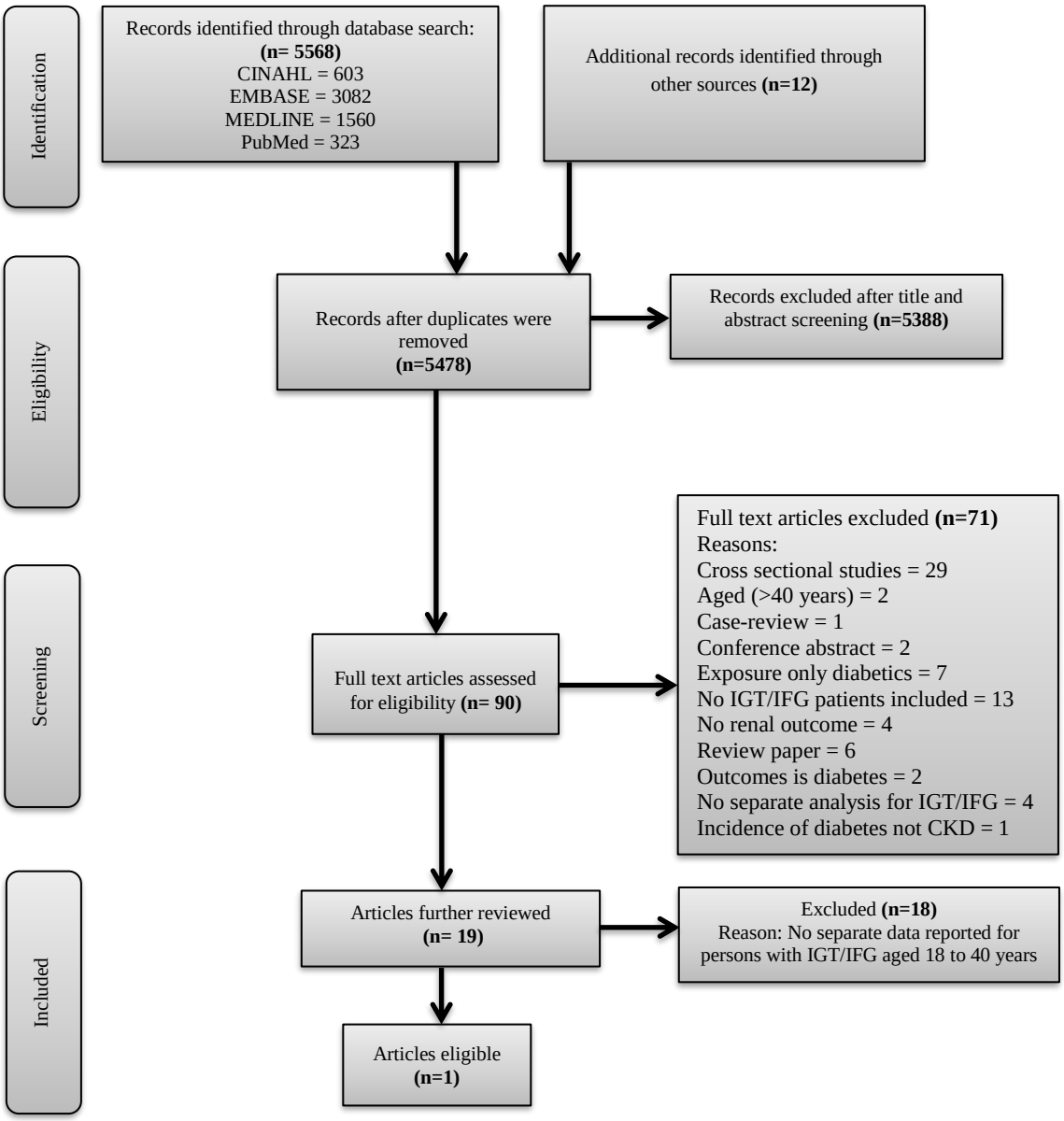


Table 5: Quality assessment of included studies: IGT/IFG compared to normoglycaemia

Study	Selection of participants	Adequate description of study population	Validated method to ascertain exposure	Validated method to confirm outcome	Adequate follow-up	Completeness of follow-up (attrition)	Analysis control for confounding	Sample size calculation	Analytical methods appropriate	Adjustment for confounders
Nelson et al (1996)(102)	YES	YES	YES	YES	YES	YES	YES	Unclear	YES	None
Fox et al (2005)(5)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	Age, sex, baseline GFR, SBP, hypertension treatment, smoking, BMI, total & HDL cholesterol, MI, congestive heart failure
Meigs et al (2002)(162)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	Age, SBP, BMI, smoking, ACE inhibitor, total cholesterol, HDL, triglyceride, hypertensive drugs.
Nelson et al (1999)(163)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	None
Nelson et al (1989)(164)	YES	YES	YES	YES	YES	YES	YES	NO	YES	Age, sex, BP
Yokoyama et al (2009)(165)	YES	YES	YES	YES	NO	Unclear	YES	NO	YES	None
Tozawa et al (2007)(166)	YES	YES	YES	YES	YES	NO	YES	NO	YES	Age, sex, current cigarette smoking, alcohol drinking habit
Nelson et al (1993)(167)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	None
Rashidi et al (2007)(168)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	None
Kitiyakara et al (2007)(169)	YES	YES	YES	YES	YES	NO	YES	NO	YES	Age, sex and smoking status
Sun et al (2010)(6)	YES	YES	YES	YES	YES	YES	YES	NO	YES	Age, sex, check-up centers and current smoking
Yang et al (2012)(170)	YES	YES	YES	YES	YES	NO	YES	NO	YES	Age, sex, BMI, serum level, total cholesterol, BP, triglyceride, HDL, waist circumference
Kovacs et al (2013)(171)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	No adjustments
Watanabe et al (2010)(7)	YES	YES	YES	YES	YES	NO	YES	NO	YES	Sex and age
Ryu et al (2009)(172)	YES	YES	YES	YES	Unclear	NO	YES	NO	YES	Age, baseline GFR, glutamyltranspeptide, uric acid, triglyceride, HDL cholesterol, BP, obesity
Tohidi et al (2012)(173)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	BMI, total cholesterol, SBP
Jee et al (2005)(174)	YES	YES	YES	YES	YES	NO	YES	NO	YES	None
Carson et al (2015) (175)	YES	YES	YES	YES	YES	YES	YES	NO	YES	Age, race, sex, education
Nand et al (2015) (176)	YES	YES	YES	YES	Unclear	Unclear	NO	NO	YES	None
Halbesma et al (2008) (177)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	Urea excretion, cholesterol, waist circumference, BP, glucose
Schottker et al (2013)(178)	YES	YES	YES	YES	YES	YES	YES	NO	YES	BMI, BP, cholesterol, antihypertensive drug, statins, smoking, history of CVD
Bonnet et al (2006) (179)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	Age, ACE inhibitors, smoking, fibrinogen level
Lucove et al (2008) (180)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	age, sex, study centre, education, and smoking
Ninomiya et al (2006) (181)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	Age, sex, baseline GFR, proteinuria, serum albumin, cholesterol, haemoglobin, alcohol, tobacco use, hyperinsulinemia

Table 6: Quality assessment of included studies: IGT/IFG compared to T2DM

Study	Selection of participants	Adequate description of study population	Validated method to ascertain exposure	Validated method to confirm outcome	Adequate follow-up	Completeness of follow-up (attrition)	Analysis control for confounding	Sample size calculation	Analytical methods appropriate	Adjustment for confounders
Kim et al (2010)(182)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	Age and sex
Iseki et al (2004)(183)	YES	YES	YES	YES	YES	Unclear	YES	NO	YES	Age, sex, baseline GFR, SBP,DBP,BMI, total cholesterol, triglyceride, serum creatinine, hematuria and proteinuria

[4]GFR, glomerular filtration rate; SBP, systolic blood pressure; BMI, body mass index; HDL, high density lipoprotein; MI, myocardial infarction; BP, blood pressure; DBP, diastolic blood pressure; FPG, fasting plasma glucose

3.6.2 Study characteristics

The characteristics of the 25 cohort and one case-control studies are summarised in Appendix 9 and Appendix 10. Briefly, one case-control and 25 cohort studies meeting the inclusion criteria were identified following the sensitivity search. One reported separate data in persons aged 18 to 40 years with IGR compared to T2DM and this was identified in the original search:

- Incidence of CKD in persons aged (18 to 40 years) with IGR compared to T2DM

Kim and colleagues, (2010) (182), reported risk of CKD in young adults aged 18 to 40 years with IGR compared to T2DM. This cohort study followed 2,666 Pima Indian youth aged (≤20 years) with IGR and type 2 diabetes (T2DM) during a follow-up period of 25.2 year for the development of macro-albuminuria, defined as an albumin creatinine ratio (ACR) of ≥300mg/g. The incidence of macro-albuminuria was 1.3 (new cases of macro-albuminuria per 1000 person-years) with a total of 28 cases in 21,830 (person-years) of follow-up in subjects with IGR or 0.13% developing macro-albuminuria each year compared to 2.4% in patients with T2DM.

3.7 Discussion

This systematic review aimed at exploring the risk of CKD in persons aged 18 to 40 years with IGR compared to those with normoglycaemia or T2DM. Nineteen cohort studies were identified for potential inclusion where some participants were aged 18 to 40 years with IGT or IFG compared to normoglycaemia or T2DM. No case-control studies were identified reporting incident cases of CKD person aged 18 to 40 years exposed to IGT/IFG compared to normoglycaemia or T2DM. Only one study reported risk estimates of CKD (macro-albuminuria ≥300mg/g) development exclusively in young adult aged 18 to 40 years with IGR compared to a similar population with T2DM. With IGR the annual incidence rate of CKD was 0.13% compared to 2.4% with T2DM. The remaining 18 studies did not report separate result in

persons aged 18 to 40 years with IGT/IFG and risk of CKD but results were reported for the population as a whole, consequently pooled estimates of CKD and a meta-analysis could not be performed because risk estimates of CKD in persons with IGT/IFG compared to normoglycaemia/T2DM were not stratified by age and reported separately. Therefore, quantification of risk was not possible in this age group.

In a systematic review of 9 population-based cohort studies, Echouffo-Tcheugui and colleagues (184) examined the effect of prediabetes defined as (IGT or IFG) and risk of CKD in a total population of 185,452 participants aged (≥18 years) with 835,146 person-years of follow-up. Incident diagnosis of CKD (the outcome of interest) was determined by any of the following definitions (micro-albuminuria, albuminuria, or proteinuria) and/or a decreased in estimated glomerular filtration rate (eGFR <60ml/min/1.73m^2). A random effects meta-analysis was used to calculate pooled risk ratios of incident CKD. The summary effect estimate showed an 11% (95% CI, 1.02 to 1.21) increased risk of CKD in individuals with prediabetes. The researchers concluded that the long term association of prediabetes and CKD may be stronger than that found in the study because of short and intermediate follow-up period of studies included in this systematic review. Additionally, the effect of prediabetes and future risk of CKD were not assessed by age group, therefore the true extent of CKD risk associated with prediabetes stratified by age group remains to be clarified.

3.8 Strength and limitations

3.8.1 Strengths

This review was not limited to the English language or geographical area and a broad range of markers were used to ascertain CKD. Furthermore, a comprehensive literature search was conducted on a topic where to the best of our knowledge, no systematic review evaluated the risk

of CKD in young adults aged 18 to 40 years with IGT/IFG compared to normoglycaemia or T2DM. Additionally, only cohort and case-control studies with incident CKD were reviewed for potential inclusion, cross-sectional studies were excluded because they do not distinguish between IGR diagnosed after CKD and IGR diagnosed before CKD.

3.8.2 Limitations

Only one study provided risk estimates of CKD in persons aged 18 to 40 years with IGR. Sufficient studies were not available to conduct a meta-analysis therefore a more generalisable and precise estimate of CKD could not be presented. Furthermore, results of this study should be interpreted with caution because of the small sample size and the study population (Pima Indians). Additionally, quantification of CKD risk was problematic because of this narrow age group, as risk of CKD in IGR in separate age groups was not available in existing literature. Researchers will be contacted requesting data to perform a sub-group analysis in the 18 to 40 age group.

3.9 Conclusion

Results of this systematic review demonstrate that evidence for the risk of CKD in young adults aged 18 to 40 years with IGR is lacking. Further research is needed to estimate the incidence of CKD in this cohort of individuals. To bridge this gap in evidence large epidemiological databases may be examined to quantify the risk of CKD in young adult aged 18 to 40 years with IGR compared to those with normoglycaemia. Data from these databases may potentially inform a prognostic study which may be useful in understanding the course and factors associated with CKD development. Finally, results may emphasise the importance of identifying individuals with IGR earlier and implement interventions to prevent or delay the development of CKD.

CHAPTER FOUR: METHODS
INCIDENCE AND PREVALENCE

4.1 Introduction

This chapter sets out to describe the method used to determine the risk of chronic kidney disease (CKD) in young adults aged 18 to 40 years with impaired glucose regulation (IGR). The chapter begins by exploring the incidence and period prevalence of IGR in young adults followed by the incidence of CKD in a similar population with IGR compared to those with normoglycaemia. Next, the incidence of each CKD category (stages 1 – 2 /3 – 5) was determined separately in the IGR cohort. It continues by providing a brief introduction to the Health Improvement Network (THIN) Database followed by describing the criteria used to identify eligible patients, the outcome of interest, estimate baseline clinical measurements, identify covariates. The chapter ends by describing the appropriate method related to the analysis of the data.

4.2 Study design

This was a population based, matched retrospective, open cohort study using data from the THIN database. An open cohort design was used because it is dynamic, meaning patients aged 18 to 40 years with IGR can enter or leave throughout the study period. Patients are continually added when they are diagnosed, replacing those who were lost to follow-up or died. A matched cohort design was chosen so that the comparison group (individuals with no glycaemic abnormality) is similar to the index group (individuals with IGR) with respect to age, gender, registered practice and date of registration.

4.3 Data source

Routinely collected electronic health records of patients registered with a practice contributing data to THIN were extracted and included in the analyses. THIN contains anonymised electronic patient records from over 600 general practices from across England, Wales, Scotland and Northern Ireland. As of September 2014 the database, holds a total of 13 million patients of whom 3.6 million are actively registered with a general practice equivalent to more than 85 million person years of computerised data covering approximately 6% of the UK population (185) (http://www.epic-uk.org/our-data/statistics.shtml). The information recorded in THIN include demographic details, diagnoses, prescriptions, socio-economic data, free text comments and additional health records such as lifestyle habits (smoking and alcohol consumption). General practices contributing data to THIN are made up of practices using Vision software which codes clinical data using the 5-byte Read code clinical classification (131) and store drug prescriptions using their generic names which can be linked to the British National Formulary (133).

4.4 Practice eligibility criteria

Practices were eligible to contribute patients to the study cohort from one year after installation of the Vision software and have an Acceptable Mortality Rate (AMR) date (a marker for data quality in the THIN database). This ensured that practices were using their system adequately and important events were not missed (134). Practices continue to contribute patients to the study cohort until the date of the last data collection (maximum February 2015).

4.5 Patient eligibility criteria

Patients were eligible for inclusion in the study from the latest of the following four dates: Vision installation date, study start date (January 2000), patient registration date plus one year

(to ensure patients did not have the outcome of interest (CKD) prior to entry) and were aged 18 years. Patients were followed up until the earliest of the following dates: aged 41 years, patient transferred out of the practice, patient died and study end date (February 2015), last data collection date. These dates were the patients start and end dates (their observation period).

4.5.1 Exclusion criteria

Patients who had no or inconsistent registration and re-registration dates and patients who were temporarily registered were excluded. Patients with an observation period of less than 90 days were also excluded from the study cohort.

4.6 Exposure (IGR case definition)

Patients with incident IGR within their observation period were eligible for inclusion as cases in the exposure cohort. The IGR index date was identified using the earliest of the following events:

1. A Read coded diagnosis of IGR (IGT + IFG or Pre-diabetes) – Refer to Appendix 11
2. Reported measurements of:
    a. HbA1c in the range 42 – 47 mmol/mol (65, 186)
    b. Fasting blood glucose (FBG) in the range 6.1 – 6.9 mmol/L (65, 67)
    c. Oral Glucose Tolerance Test (OGTT) blood glucose in the range ≥7.8 ≤11.1 mmol/L (65, 67)

4.6.1 Exclusion criteria

a. Patients with an index date prior to their observation period
b. Patients with diabetes (type1 or type 2) prior to the index date
c. Patient observation period ended within 90 days of the index date

d. Patient diagnosed with diabetes (type 1 or type 2) within 30 days of the index date

e. Patients with CKD prior to their index date

4.7 Outcome (CKD case definition)

Patients with incident CKD stage 1 – 5 were identified using the earliest of the following events:

1. A Read coded diagnosis of CKD (1 – 2/ 3 – 5) – (Appendix 11) or
2. A single estimated glomerular filtration rate (eGFR) of (≥60 ml/min/1.73m^2 [CKD stage 1 – 2]) with additional evidence of kidney disease which include the following:
    a. A single albumin creatinine ratio (ACR) of (≥3mg/mmol) (9)
    b. A single protein creatinine ratio (PCR) of (50mg/mmol) (9)
    c. A single urinary albumin/protein loss of 0.5g/day (9)
3. eGFR<60 ml/min/1.73m^2 (CKD stage 3 – 5)

4.8 Covariates

Covariates (risk modifiers) were selected based on their biological plausibility and previously published epidemiological evidence (54, 58). These covariates were age hypertension, heart failure (HF), atrial fibrillation (AF), cardiovascular disease (CVD), and non-steroidal anti-inflammatory drugs (NSAIDS). The latter was described as the date of the latest prescription prior to the index date. All the other covariates were recorded before the index date.

4.9 Matching criteria

Eligible IGR cases were matched on practice, gender and age (within 2 years) and IGR case index date. For each IGR case three unexposed controls (normoglycaemia) were selected.

Matched patients had to remain registered for at least one year prior to the index date of the IGR case and not have IGR or CKD on or before the index date.

4.9.1 Exclusion

1. Potential matches were excluded if:
    a. Their observation period ended within 90 days of the index date
    b. Diagnosed with diabetes (type 1/type 2) within 30 days of the index date

4.10 Follow-up period

1. IGR patients were followed up until the earliest of the following dates:
    a. Patient died
    b. Left practice
    c. Diagnosis of CKD
    d. Practice last data collection

4.11 Extraction by Read code

A Read code list was compiled to facilitate data extraction which was sufficiently accurate and valid for this study. This list was derived from the quality and outcomes framework (QOF) version (33.0) 2015 rulesets (http://www.hscic.gov.uk/qofbrv33). The Read code list was checked and approved by Professor Tom Marshall, Dr Paramjit Gill and Dr Ronan Ryan. A full list of Read codes for each study variable included in this study and a description of what each Read code relates to, are presented in Appendix 11.

4.12 Missing data

Several studies (187, 188) have shown that databases such as THIN are known to contain some missing data. However, excluding missing data (for e.g., only including complete case analysis) may result in unstable and inaccurate analysis (189). A challenge in THIN is that data are generally captured for clinical purposes and not for research which in some circumstances may result in data missing due to unobserved characteristics which are not likely to be missing at random (190). Furthermore, imputing missing values require the data to be missing at random (191). Additionally, some patients may consult infrequently and are likely to have missing data (for e.g., BMI, ethnicity) and are therefore less likely to be diagnosed with an illness. Therefore, no attempt was made to impute missing data. For categorical variable (e.g. ethnicity) and continuous variables (e.g. BMI) a separate 'missing' category was created for those with missing data.

4.13 Ethics approval

THIN database (data collection process) obtained ethical approval from the South East Multi-centre Research Ethics Committee (reference number: 07/H1102/103). The protocol for this study was reviewed and approved in June 2014 (study reference: 14-038) by the THIN Scientific Review Committee (SCR) (Appendix 12).

4.14 Statistical analysis

4.14.1 Descriptive statistics

All statistical analyses were conducted using Stata version 13.1 (Stata Corp, College Station Texas, USA). Appropriate descriptive statistics were used to summarise the covariates and matching characteristics for those exposed and unexposed to IGR. Categorical variables were investigated using (Chi-square [χ^2]) test and continuous variables were analysed using t-test.

Since, length of follow-up had a highly skewed distribution Mann-Whitney U test was used to compare groups. Statistical significance was set at 5% level (p – value = 0.05).

4.14.2 Estimating IGR incidence

The overall incidence rate of IGR amongst young adults aged 18 to 40 years was calculated by dividing the number of events (failure) that occurred by person years of follow-up (period of January 2000 to February 2015). Similarly, incidence rates by age group (18-25, 26-40), gender, ethnicity, deprivation quintile and calendar year (2000 - 2015) were also calculated. Incidence rates were expressed per 100,000 person-years of follow-up with 95% confidence interval (CI). Patients contributed person-years until the earliest of the following: developed diabetes, end date (died, left practice, practice last data collection date, earliest of age 40). STRATE command was used to examine incidence rates of IGR per 100,000 years of follow-up (192).

4.14.3 Calculating period prevalence of diagnosed IGR with mid-year study population

The period prevalence of diagnosed IGR amongst young adults aged 18 to 40 years was estimated over the study period (January 2000 to December 2014). Annual prevalence of diagnosed IGR was calculated by age group, gender and calendar year. The prevalence of IGR was estimated by including all new and pre-existing cases of IGR in each or any previous year as numerator divided by the mid-year study population in that year. The average population in the practice is calculated from the number of persons registered during the year. Patients with IGR were identified by Read code only because laboratory results recorded in their previous practice's electronic records were not routinely transferred into the electronic record at their current (THIN) practice. Using a definition of IGR based on laboratory results would therefore mean we were more likely to identify cases diagnosed after registration with a THIN practice than before. We decided that it was preferable to adopt a definition based on Read codes alone as

it would avoid introducing an ascertainment bias. The cost of this decision is that our estimates of prevalence are likely to be lower than the true prevalence, but this was felt to be a more consistent way of representing prevalence over time. Period prevalence was expressed as a percentage with a corresponding 95% confidence interval (CI). Patients were excluded if they developed diabetes during the study period, aged (>40 years), died, left practice and practice last data collection date. A combination of STSPLIT and STRATE commands in STATA were used to calculate period prevalence of IGR by calendar year.

4.14.4 Calculating incidence rate of CKD in IGR compared to those without IGR

The incidence rate ratio of CKD in patients aged 18 to 40 years for the period January 2000 to February 2015 was estimated by dividing the incidence rate of CKD in patients with IGR to those with normoglycaemia. After patients developed IGR they were assumed to be exposed from this date onwards. There was no allowance for regression from IGR back to normoglycaemia. Incidence rate was expressed as 100,000 person-years of follow-up. Adjusted and unadjusted incidence rate ratio (IRR) of CKD was calculated with 95% CI and associated p – values matched on age, sex and practice. The STIR command was used to examine the unadjusted IRR of CKD. Poisson regression (Poisson command in STATA) (193) was used to adjust for patient level covariates (age, sex, ethnicity, level of deprivation, BMI categories, CVD, HF, AF, hypertension, NSAID). Patients were censored if they developed diabetes during follow-up, died, left practice, practice last data collection date or had a CKD diagnosis.

4.14.5 Calculating incidence rate of CKD (stages 1 – 2 /3 – 5) separately in IGR

The incidence rate of each CKD category (stages 1 – 2 /3 – 5) was calculated separately. Patients were classified as CKD stage (1 – 2/3 – 5) using the first recorded Read code or first recorded laboratory data. Incidence rates were expressed per 100,000 person-years of follow-up with 95%

(CI). STRATE command was used to examine the unadjusted incidence rate of CKD 1 – 2 and CKD 3 – 5 in IGR per 100,000 person-years of follow-up. Patients who were categorised as CKD stage (1 – 2) and later progressed to CKD stage (3 – 5) were censored at the date when they progressed from CKD stage (1 – 2) to (3 – 5) during follow-up.

CHAPTER FIVE: METHODS
PREDICTORS OF CKD IN IGR

5.1 Introduction

The aim of this chapter was to determine future CKD risk in patients following a diagnosis of IGR. This was done using a Cox-proportional hazards model. The relationship between each baseline characteristic and CKD risk was assessed separately in a series of univariate analyses. This was followed by checking for a non-linear relationship between CKD risk and the continuous variables age and BMI by fitting fractional polynomials (194).

Next, the probability of remaining free of CKD was estimated up to a maximum of 15 years following IGR diagnosis using Kaplan-Meier survivor function. The proportional hazards assumption for each variable was checked using Schoenfeld residuals and log-log survival plots. A multivariate Cox proportional hazards model containing all the predictor variables was then fitted. This chapter ends by exploring the goodness of fit of the final model, estimated by using the R-squared (R^2) statistic for time-to-event models developed by Royston and colleagues (194).

5.2 Method

The study design, data source, patient and practice eligibility criteria, IGR case definition, outcome definition and prognostic variables included in this analysis have been previously discussed in chapter 4 (methods chapter: prevalence and incidence). Briefly, data from 10,561 patients with incident IGR were included in this analysis. IGR cases were defined as patients without CKD at diagnosis of IGR and the outcome of interest was time to CKD diagnosis. IGR

patients were followed up until the earliest of the following events: 1) end of the study period, death or de-registration; 2) diagnosis of CKD; 3) diagnosis of diabetes (no longer IGR). Prognostic variables were chosen based on literature and clinical importance (54, 58).

5.3 Model development

5.3.1 Univariate analysis

Each CKD risk predictor was tested separately by fitting Cox regression (195) using the *stcox* (196) command. A hazard ratio (HR) with 95% confidence interval (CI) and associated p-value were derived for each predictor. Predictors were tabulated for up to 15-years follow-up, with the aim of excluding predictors from the multivariate analysis where it was not possible to reliably estimate risk of CKD.

5.4 Handling of continuous predictor: Checking for non-linear relationships between the continuous variable and the outcome

Fractional polynomials (194) were used to check for a non-liner relationships between each of the continuous variables (BMI and age) and risk of CKD. The Stata command *mfp stcox* (197) was used to determine the best fitting functional form for each variable.

5.5 CKD free survival following IGR diagnosis

A Kaplan-Meier (K-M) survivor function (198) was used to determine probability of remaining CKD free from IGR diagnosis to a maximum of 15 years. The Kaplan-Meier function estimates probability of survival at specified time points in the presence of censored data. The *sts list* (199) command was used to retrieve a tabular version of the survival probabilities and 95% CIs. The results of the Kaplan-Meier survival estimates were reviewed to ensure that there were sufficient

number of patients included in the analysis, particularly in later years. A final decision on the length of follow-up was made using this information and clinical colleagues' advice.

5.6 Method for model checking – checking for proportional hazards assumption

The Cox model makes the assumption that the effect of the predictor variables and time to event (risk of CKD) remains constant over time. This assumption was checked for each predictor and globally using the *estat phtest detail* (200) command. It produces Chi-square [χ^2] and a p-value for each predictor and globally. Proportionality was also checked visually using log-log graphs. The *stphplot* (200) command was used to draw log-log plots for the selected covariates. Hazards can be assumed proportional if the plotted line for each group remained roughly parallel with its neighbours over time (201).

5.7 Multivariate analysis

A multivariate Cox proportional hazards model was fitted containing predictors with reliable CKD risk estimates reported in the univariate analyses to determine risk of CKD in patients with IGR for up to 15-years follow-up. Risk of CKD was reported as hazard ratios (HR) along with 95% confidence intervals (CI) and associated p-values.

5.8 Model checking – proportion of variation in the final model

To assess goodness of fit R-squared (R^2) statistics was used. The command *str2ph* is based on Royston's (202) modification of Nagelkerke's statistics for proportional hazards models (203). It estimates the proportion of variation in the data explained by the model. The R^2 ranged from 0 to 1, where 1 indicates a perfect model fit in predicting the outcome for each case.

CHAPTER SIX: RESULTS
INCIDENCE AND PREVALENCE

6.1 Introduction

This chapter presents the findings of the following analyses of the Health Improvement Network (THIN) database. The first part describes the incidence and period prevalence of impaired glucose regulation (IGR) in persons aged 18 to 40 years by age group, gender, ethnicity, area of deprivation and calendar year. After this the prevalence of IGR is compared to two published papers reporting prevalence of prediabetes and IGT/impaired fasting glucose (IFG) (Appendix 13).

The second part of the analysis consists of the unadjusted and adjusted incidence of chronic kidney disease (CKD) in persons aged 18 to 40 years with IGR compared to those with no glycaemic abnormality. The final part presents the results of the unadjusted incidence of CKD 1 – 2 and 3 – 5 separately in the IGR cohort. The chapter concludes with a description of the baseline characteristics of the matched cohort (IGR compared to normoglycaemia in persons aged 18 to 40 years) and the number of CKD cases first recoded via Read code and first recorded via biomedical data in the IGR cohort.

6.2 Incidence of IGR in persons aged 18 to 40 years

6.2.1 Overall incidence of IGR

The study population included in this analysis consists of approximately 3.5 million individuals aged 18 to 40 years contributing 14.9 million person-years (pyr) of follow-up for the period of (January 2000 to December 2014). The overall incidence rate (IR) of IGR was 72.2 cases per

100,000 pyr (95% CI, 70.9 to 73.6) of follow-up. The incidence rates per (100,000 pyr) by age group, gender, ethnicity, area of deprivation and calendar year are presented in Table 7.

6.2.2 Incidence by age group and gender

The incidence of IGR increased with age. For those aged 26 to 40 years the incidence rate was approximately 5 times higher (IR 93.8 [95% CI, 91.9 to 95.6]) per 100,000 pyr compared to individuals aged 18 to 25 years, the incidence rate was (22.1 [95% CI, 20.8 to 23.5]). The incidence rate was higher in females (79.0 [95% CI, 77.0 to 81.0]) per 100,000 pyr of follow-up as compared to males (65.8 [95% CI, 64.0 to 67.7]) per 100,000 person-years.

6.2.3 Incidence by ethnicity

The incidence of IGR in South Asians was approximately 5 times higher (IR 393.7 [95% CI, 374.5 to 413.9]) and in Black approximately 3 times higher (IR 218.1 [95% CI, 197.9 to 240.4]) than the incidence in the White ethnic group (IR 76.5 [95% CI, 74.1 to 79.0]) per 100,000 pyr. The incidence rate for patients with missing information on ethnicity was 52.7 (95% CI, 51.3 to 54.2) per 100,000 pyr.

6.2.4 Incidence by deprivation quintile

The incidence of IGR increases with increasing deprivation. The incidence rate was approximately 2.2 times higher (IR 105.5 [95% CI, 101.2 to 109.9]) in the most deprived areas (Townsend quintile 5) compared to the least deprived areas (IR 48.1 [95% CI, 45.8 to 50.6]).

6.2.5 Incidence by calendar year

Table 7, illustrates the annual incidence of IGR for the period (2000 – 2014). The incidence of IGR was approximately 8.4 times higher in 2014 (IR 137.3 [95% CI, 130.1 to 144.8]) than 2000 (IR 16.4 [95% CI, 13.0 to 20.6]). This increasing trend appears to continue throughout the 15

year period. The highest peak recorded was in 2013, the incidence rate was more than 9 times higher (IR 157.1 [95% CI, 149.8 to 164.7]) than incidence in 2000.

Table 7: Incidence of IGR by age, sex, deprivation score, ethnicity and calendar year in THIN database from January 2000 to December 2014

	New cases of IGR	Total population at risk	Person-years	Incidence rate per 100,000 person-years	95% (CI)
Total	10,776	3,456,259	14,915,070	72.2	(70.9, 73.6)
Age group					
18-25	991	1,464,328	4,479,930	22.1	(20.8, 23.5)
26-40	9785	2,579,647	10,435,140	93.8	(91.9, 95.6)
Gender					
Male	5019	2,012,730	7,621,760	65.8	(64.0, 67.7)
Female	5759	2,053,522	7,293,310	79.0	(77.0, 81.0)
Ethnicity					
White	3727	1,396,260	4,874,090	76.5	(74.1, 79.0)
Black	405	62,697	185,700	218.1	(197.9, 240.4)
Asian	1535	125,643	389,870	393.7	(374.5, 413.9)
Chinese	32	16,909	38,530	83.1	(58.7, 117.4)
Mixed	63	24,352	67,370	93.2	(73.1, 119.7)
Others	132	37,024	103,750	127.2	(107.3, 150.9)
Unknown	4882	2,403,367	9,255,760	52.7	(51.3, 54.2)
Townsend Quintile					
(Least deprived) 1	1540	811,852	3,202,050	48.1	(45.8, 50.6)
2	1603	722,848	2,811,570	57.0	(54.3, 59.9)
3	2216	834,840	3,092,000	71.7	(68.7, 74.7)
4	2520	830,643	2,941,250	85.7	(82.4, 89.1)
(Most deprived) 5	2283	617,112	2,164,970	105.5	(101.2, 109.9)
Unknown	614	248,957	703,230	87.3	(80.7, 94.5)
Incidence by calendar year					
2000	73	597,233	445,640	16.4	(13.0, 20.6)
2001	105	741,360	626,240	16.8	(13.8, 20.3)
2002	210	928,062	770,200	27.3	(23.8, 31.2)
2003	338	1,012,844	898,260	37.6	(33.8, 41.9)
2004	425	1,056,516	945,550	44.9	(40.9, 49.4)
2005	538	1,141,176	1,012,820	53.1	(48.8, 57.8)
2006	519	1,182,937	1,054,750	49.2	(45.2, 53.6)
2007	635	1,222,252	1,080,410	58.8	(54.4, 63.5)
2008	651	1,279,799	1,126,940	57.8	(53.5, 62.4)
2009	849	1,310,172	1,163,760	73.0	(68.2, 78.0)
2010	915	1,292,454	1,146,900	79.8	(74.8, 85.1)
2011	1020	1,270,332	1,134,160	89.9	(84.6, 95.6)
2012	1297	1,275,445	1,132,720	114.5	(108.4, 120.9)
2013	1706	1,246,200	1,085,980	157.1	(149.8, 164.7)
2014	1351	1,143,592	984,250	137.3	(130.1, 144.8)

Abbreviations: IGR, impaired glucose regulation; CI, confidence interval, PYR; person-years

6.3 Read coded period prevalence of diagnosed IGR in persons aged 18 to 40 years based on mid-year population estimates

A total of approximately 3.5 million individuals aged 18 to 40 years were included in this analysis. The annual prevalence of diagnosed IGR by calendar year, gender and age group for the period (January 2000 to December 2014), are presented in Tables (8 - 10).

6.3.1 Period prevalence of diagnosed IGR by calendar year based on mid-year study population for each year (2000 to 2014)

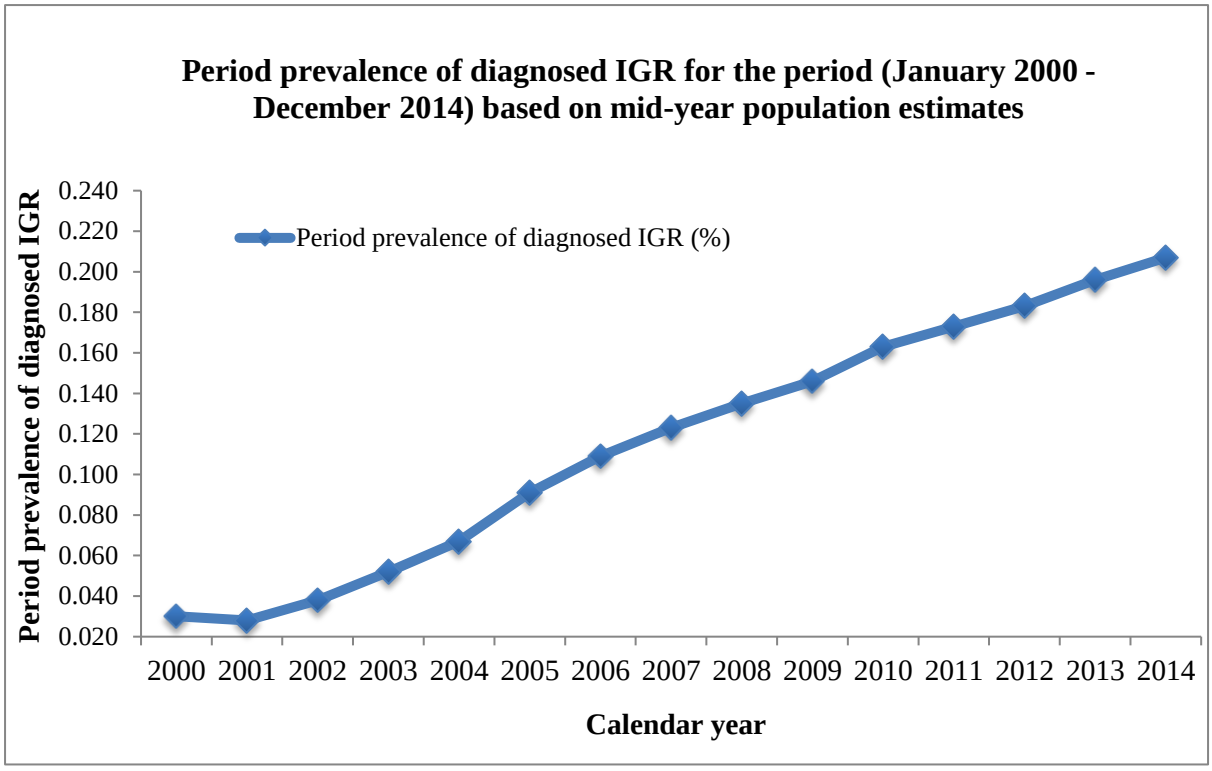
The prevalence of diagnosed IGR increased 7 fold, from 0.030% (95% CI, 0.025% to 0.036%) in 2000 to 0.207% (95% CI, 0.198% to 0.216%) in 2014 (Table 8). The annual prevalence consistently increased throughout the 15 year period (Figure 9).

Table 8: Period prevalence of diagnosed IGR in persons aged 18 to 40 years by calendar year, 2000 to 2014 based on mid-year population estimates

Year	Number of patients with IGR	Persons (Mid-year population)	Prevalence (%)	95% (CI)
2000	131	434,340	0.030	(0.025, 0.036)
2001	175	626,262	0.028	(0.024, 0.032)
2002	294	769,503	0.038	(0.034, 0.043)
2003	476	907,332	0.052	(0.048, 0.057)
2004	634	946,531	0.067	(0.062, 0.072)
2005	926	1,022,355	0.091	(0.085, 0.097)
2006	1149	1,050,507	0.109	(0.103, 0.116)
2007	1333	1,082,085	0.123	(0.117, 0.130)
2008	1519	1,127,671	0.135	(0.128. 0.142)
2009	1706	1,171,734	0.146	(0.139, 0.153)
2010	1865	1,144,575	0.163	(0.156, 0.171)
2011	1968	1,136,167	0.173	(0.166, 0.181)
2012	2079	1,134,967	0.183	(0.175, 0.191)
2013	2152	1,099,471	0.196	(0.188, 0.204)
2014	2052	993,259	0.207	(0.198, 0.216)
Average annual prevalence			0.116	

Abbreviation: IGR, impaired glucose regulation; CI, confidence interval

Figure 9: Trend in the prevalence of diagnosed IGR in persons aged 18 to 40 years for the period (January 2000 – December 2014) based on mid-year population estimates



6.3.2 Period prevalence of diagnosed IGR by age group based on mid-year study population for each year (2000 to 2014)

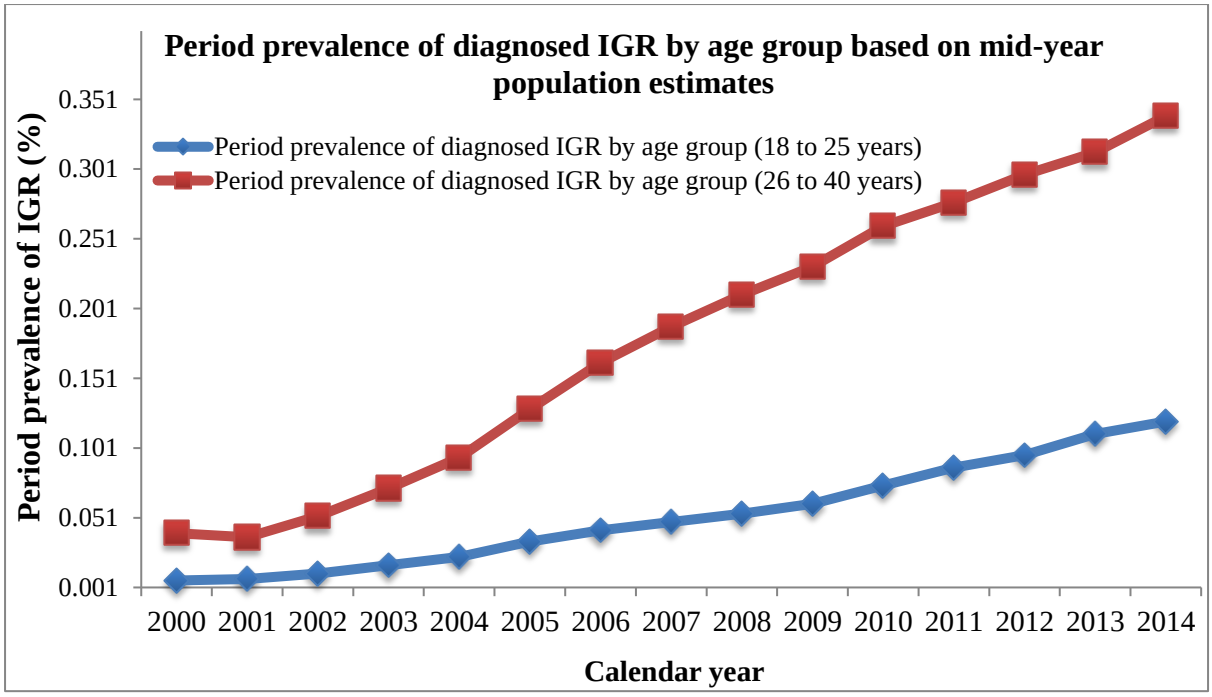
The prevalence of diagnosed IGR in the age group 18 to 25 years was lower than in those aged 26 to 40 years over the 15 year calendar period (January 2000 to December 2014) (Figure 10). Prevalence increased in both age groups from 0.040% (95% CI, 0.033% to 0.047%) in 2000 to 0.339% (95% CI, 0.321% to 0.358%) in 2014 in those aged 26 to 40 years and from 0.006% (95% CI, 0.003% to 0.012%) in 2000 to 0.120% (95% CI, 0.112% to 0.129%) in 2014 in those aged 18 to 25 years. The annual prevalence of IGR by age group is shown in Table 9. Prevalence of IGR in this study was compared with Mainous and colleagues (76) and the NHS Diabetic Prevention Programme (NHS-DPP) Non-diabetic hyperglycaemia (77) (Appendix 13).

Table 9: Prevalence of diagnosed IGR in persons aged 18 to 40 years by age group based on mid-year population estimates

	Aged (18 to 25 years)				Aged (26 to 40 years)				Ratio aged (26-40:18-25)
Year	Number of patients with IGR	Persons (Mid-year population)	Prevalence (%)	95% (CI)	Number of patients with IGR	Persons (Mid-year population)	Prevalence (%)	95% (CI)	
2000	8	124,952	0.006	(0.003, 0.012)	123	309,388	0.040	(0.033, 0.047)	6.7
2001	14	192,422	0.007	(0.004, 0.012)	161	433,840	0.037	(0.032, 0.043)	5.3
2002	28	253,871	0.011	(0.008, 0.016)	266	515,632	0.052	(0.046, 0.053)	4.7
2003	53	320,435	0.017	(0.013, 0.022)	423	586,897	0.072	(0.065, 0.079)	4.2
2004	84	361,963	0.023	(0.019, 0.029)	550	584,568	0.094	(0.087, 0.103)	4.1
2005	140	415,256	0.034	(0.029, 0.040)	786	607,099	0.129	(0.121, 0.139)	3.8
2006	190	457,456	0.042	(0.036, 0.048)	959	593,051	0.162	(0.152, 0.172)	3.9
2007	239	499,927	0.048	(0.042, 0.054)	1094	582,158	0.188	(0.177, 0.200)	3.9
2008	294	546,131	0.054	(0.048, 0.060)	1225	581,540	0.211	(0.199, 0.223)	3.9
2009	358	587,458	0.061	(0.055, 0.068)	1348	584,276	0.231	(0.219, 0.243)	3.8
2010	440	597,126	0.074	(0.067, 0.080)	1425	547,449	0.260	(0.247, 0.274)	3.5
2011	541	620,253	0.087	(0.080, 0.095)	1427	515,923	0.277	(0.263, 0.291)	3.2
2012	618	642,231	0.096	(0.089, 0.104)	1461	492,736	0.297	(0.282, 0.312)	3.1
2013	712	639,789	0.111	(0.103, 0.120)	1440	459,682	0.313	(0.297, 0.330)	2.8
2014	724	601,699	0.120	(0.112, 0.129)	1328	391,560	0.339	(0.321, 0.358)	2.8
Average annual prevalence			0.053				0.180	Average annual prevalence ratio aged (26-40:18-25)	4.0

Abbreviation: IGR, impaired glucose regulation; CI, confidence interval

Figure 10: Annual prevalence of diagnosed IGR by age group based on mid-year population estimates



6.3.3 Period prevalence of diagnosed IGR by gender based on mid-year study population for each year (2000 to 2014)

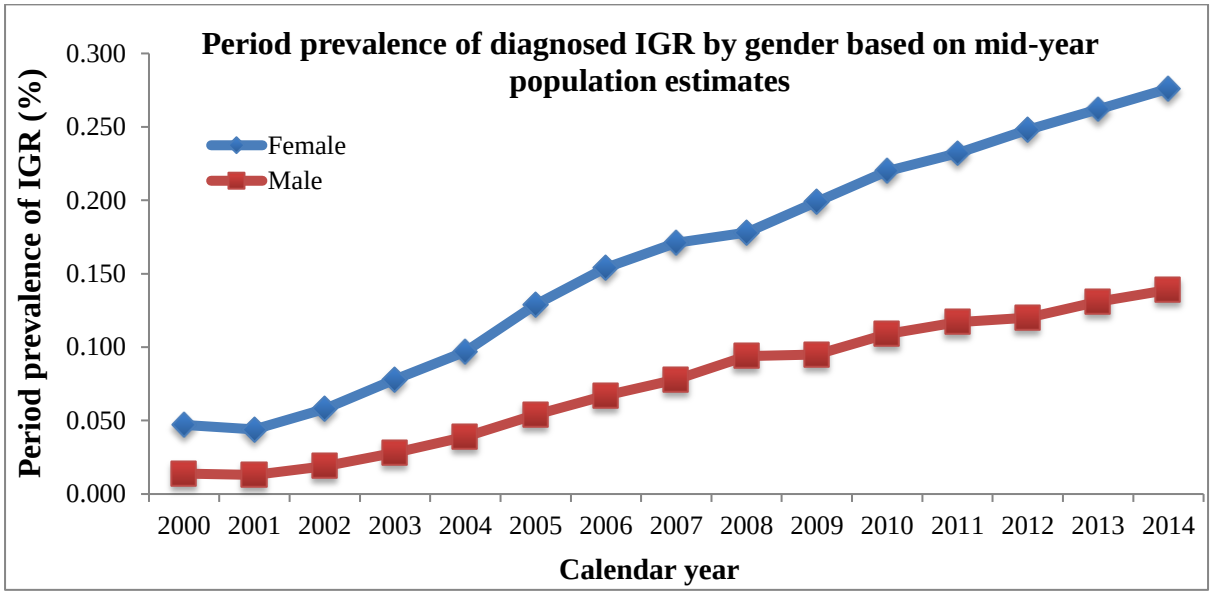
Figure 11 shows the trend in the prevalence of diagnosed IGR by gender for the study period (January 2000 to December 2014). Prevalence in females was consistently higher than males throughout the 15 year period. Prevalence increased from 0.014% (95% CI, 0.010% to 0.019%) in 2000 to 0.139% (95% CI, 0.129% to 0.150%) in 2014 in males and from 0.047% (95% CI, 0.038% to 0.057%) in 2000 to 0.276% (95% CI, 0.261% to 0.291%) in 2014 in females. The annual prevalence of diagnosed IGR by gender is shown in Table 10.

Table 10: Prevalence of diagnosed IGR in persons aged 18 to 40 years by gender based on mid-year population estimates

	Female				Male				Ratio F:M
Year	Number of patients with IGR	Persons (Mid-year population)	Prevalence (%)	95% (CI)	Number of patients with IGR	Persons (Mid-year population)	Prevalence (%)	95% (CI)	
2000	100	212,602	0.047	(0.038, 0.057)	31	221,738	0.014	(0.010, 0.019)	3.4
2001	134	305,873	0.044	(0.037, 0.052)	41	320,389	0.013	(0.009, 0.017)	3.4
2002	219	374,854	0.058	(0.051, 0.067)	75	394,649	0.019	(0.015, 0.024)	3.1
2003	346	441,593	0.078	(0.070, 0.087)	130	465,739	0.028	(0.024, 0.034)	2.8
2004	446	460,024	0.097	(0.088, 0.106)	188	486,507	0.039	(0.034, 0.045)	2.5
2005	642	497,235	0.129	(0.119, 0.139)	284	525,120	0.054	(0.048, 0.061)	2.4
2006	789	510,950	0.154	(0.143, 0.165)	360	539,557	0.067	(0.060, 0.074)	2.3
2007	898	526,358	0.171	(0.160, 0.182)	435	555,727	0.078	(0.071, 0.086)	2.2
2008	977	550,137	0.178	(0.167, 0.189)	542	577,534	0.094	(0.086, 0.102)	1.9
2009	1138	572,334	0.199	(0.187, 0.211)	568	599,400	0.095	(0.088, 0.104)	2.1
2010	1230	559,698	0.220	(0.208,0.232)	635	584,877	0.109	(0.100, 0.117)	2.0
2011	1291	557,500	0.232	(0.219, 0.245)	677	578,676	0.117	(0.109, 0.126)	2.0
2012	1385	557,428	0.248	(0.235, 0.262)	694	577,539	0.120	(0.112, 0.130)	2.1
2013	1421	542,829	0.262	(0.248, 0.276)	731	556,642	0.131	(0.121, 0.141)	2.0
2014	1352	490,146	0.276	(0.261, 0.291)	700	503,113	0.139	(0.129, 0.150)	2.0
Average annual prevalence			0.160				0.074	Average annual F:M ratio	2.4

Abbreviation: IGR, impaired glucose regulation; CI, confidence interval

Figure 11: Annual prevalence of diagnosed IGR by gender based on mid-year population estimates



6.4 Incidence rate of CKD in persons aged 18 to 40 years with IGR compared to those with normoglycaemia

The study population included in this analysis contained approximately 41,000 individuals aged 18 to 40 years for the period of (January 2000 to February 2015), matched on age, sex and general practice. The unadjusted incidence of CKD per 100,000 person-years (pyr) of follow-up in persons with IGR was 4 times higher [IRR 4.0, 95% CI, 3.2 to 5.1, p<0.001] compared to individuals with normoglycaemia. After adjustment for age, sex, ethnic group, deprivation quintile, BMI categories, cardiovascular disease, heart failure, atrial fibrillation, hypertension and NSAID, incidence of CKD was attenuated but still showed a significant association. The incidence was reduced to approximately 3 times higher [IRR 2.6, 95% CI, 2.0 to 3.4, p<0.001] in IGR than the normoglycaemia cohort. The unadjusted and adjusted incidence rates of CKD in patients with IGR compared to those with normoglycaemia are given in Table 11. The full model is presented in Appendix 14.

Table 11: Incidence of CKD in IGR compared to normoglycaemia

	Exposed	Unexposed	95% CI	P-Value
CKD	182	126		
Total population at risk	10,561	29,531		
Person-years of follow-up	24364.5	67351.1		
Unadjusted IRR*		4.0	(3.2, 5.1)	<0.001
Adjusted IRR**		2.6	(2.0, 3.4)	<0.001

* Matched on age, sex and practice

**Adjusted for age, sex, ethnicity, level of deprivation, BMI category, cardiovascular disease, heart failure, atrial fibrillation, hypertension and NSAID.

6.5 Incidence rate of CKD (stages 1 – 2 /3 – 5) in persons aged 18 to 40 years with IGR

The study population included in this analysis consists of 10,561 individuals aged 18 to 40 years with IGR and without CKD at baseline for the period of (January 2000 to February 2015). The unadjusted incidence of CKD stage (3 – 5) per 100,000 person-years (pyr) of follow-up in persons with IGR was approximately 4 times higher (IR 106.4 [95% CI, 71.9 to 157.5]) than the incidence of CKD stage (1 – 2) (IR 27.9 [95% CI, 23.9 to 32.7]). The unadjusted incidence rates of CKD (1 – 2) and CKD (3 – 5) in the IGR cohort are given in Table 12. The number of outcomes for each CKD category (1 – 2 and 3 – 5) in the IGR cohort is provided in Table 12a.

Table 12: Incidence of CKD stages (1 – 2/3 – 5) in the IGR cohort

	Exposed	95% CI
CKD	182	
Total population at risk	10,561	
Person-years of follow-up	24364.5	
Unadjusted IR per 100,000 pyr		
CKD stage (1 – 2)	27.9	(23.9, 32.7)
CKD stage (3 – 5)	106.4	(71.9, 157.5)

Abbreviations: CKD, chronic kidney disease; CI, confidence interval; PYR, person-years

Table 12a: Number of outcomes for each CKD category (1 – 2/3 – 5) in the IGR cohort

CKD category	**1-2**	**3-5**	**Total**
CKD outcome recorded	157	25	182

Abbreviation: CKD, chronic kidney disease

6.6 Characteristics of matched cohort study population

The study population consists of 10,561 patients with impaired glucose regulation (IGR) cases and 29,531 controls matched on age, gender, IGR case index date and practice (without glycaemic abnormality). The median age at index date was 35 years [interquartile range (IQR) 31-38 years] in both cases and controls. Fifty three percent of the cases and 53.5% of the controls were female. The median follow-up period for cases and controls was approximately 2 years [IQR: 0.81-3.16] and [IQR: 0.79-3.14] respectively. Patients who progressed to diabetes during follow-up were censored. There were statistically significant differences in the proportion of cases and controls who were South Asians (14.0% cases, 6.4% controls; White 34.5% cases and 38.3% controls). Information on ethnicity was missing for 4,817 (45.6%) in cases and 14,886 (50.5%) in control. Cases and controls were more likely to be living in areas of greater deprivation, 2,232 (21.1%) and 5,020 (17.0%) of cases and controls living in quintile 5 respectively. Cases and controls (22.9% and 23.8%) respectively, were more likely to have body mass index (BMI) of (25 to 29.9 kg/m^2) and less likely to be unknown, cases (9.5%) and controls (21.2%). A total of 190 (1.8%) of cardiovascular disease (CVD) were identified in cases and 274 (0.9%) in controls at index date.

Similarly, heart failure (HF) and atrial fibrillation (AF) were identified in 29 cases (0.3%), 21 control (0.07%) and 17 cases (0.16%) and 21 controls (0.07%) at index date respectively. More than 10% (1,112) of cases were diagnosed with hypertension at index date with only 548 (1.9%)

in controls. Almost twice as many cases 200 (1.9%) were prescribed a non-steroidal anti-inflammatory drug (NSAID) at baseline with 120 (0.4%) in controls. During follow-up 1262 (11.9%) cases developed T2DM, only 124 (0.4%) were identified in controls. During follow-up, 182 (1.7%) cases of CKD were observed in patients with IGR compared to 126 (0.4%) in the control group, patients who progressed to diabetes and subsequently developed CKD were censored. 429 (4.1%) of the 10,561 patients with IGR who progressed to diabetes during follow-up developed CKD. The vast majority of cases (81%) and controls (80%) were identified from practices across England followed by Wales, 9.2% and 9.3% of cases and controls respectively. Descriptive characteristics of the study population are summarised in Table 13.

6.6.1 Number of CKD cases recorded via Read code and laboratory data

Table 13a provides information about the number of CKD cases first recoded via Read code and first recorded via laboratory data in the IGR cohort. Patients were either identified by a Read code diagnosis or laboratory measurements indicating CKD. These two measures were also used to categorise CKD type. Laboratory measurements indicating CKD followed the definitions of CKD shown in Table 1. However, due to data limitations, it is not clear whether the laboratory measurements refer to ACR, GFR, or a combination of these. Data were assigned and categorised as CKD type (1 – 2 or 3 – 5) by Dr Ronan Ryan following data extraction.

Table 13a: Cases of CKD first recorded by Read code and first recorded by laboratory data in the IGR cohort

CKD results	**Read code**	**Laboratory data**	**Total**
CKD cases recorded	26	156	182

Abbreviation: CKD, chronic kidney disease

6.6.2 Number of IGR cases recorded via Read code and laboratory measurements

Table 13b provides information about the number of cases of IGR recorded via Read code and laboratory measurements derived from fasting blood glucose, glycosylated haemoglobin and oral glucose tolerance test.

Table 13b: Cases of IGR recorded by Readcode and laboratory measurements

IGR case definition	**Readcode**	**FBG**	**HbA1c**	**OGTT**	**Total**
IGR cases recorded	2,245	3,483	4,778	55	10,561

Abbreviations: IGR, impaired glucose tolerance; FBG, fasting blood glucose; HbA1c; glycosylated haemoglobin; OGTT; oral glucose tolerance test

6.6.3 Number of IGR cases recorded via laboratory measurements by calendar year

Table 13c provides information about the number of cases of IGR recorded via laboratory measurements by calendar year for the study period January 2000 to December 2014.

Table 13c: Cases of IGR recorded via laboratory tests by calendar year (January 2000 to December 2014)

Year	**IGR cases recorded by laboratory measurements**
2000	57
2001	82
2002	164
2003	219
2004	316
2005	373
2006	329
2007	448
2008	446
2009	596
2010	651
2011	769
2012	1074
2013	1486
2014	1176

Abbreviation: IGR, impaired glucose regulation

Table 13: Characteristics of the matched cohort (IGR compared to normoglycaemia) †

	IGR [N=10,561 (26%)]		Normoglycaemia [N=29,531(74%)]		*P*-Value
Age, median [IQR]	35	[31,38]	35	[31,38]	0.015††
Female, No. (%)	5,659	(53.6)	15,795	(53.5)	0.863†
Years of follow-up, median [IQR]	1.68	[0.81, 3.16]	1.67	[0.79, 3.14]	0.684†††
Ethnicity, No. (%)					
White	3,644	(34.5)	11,305	(38.3)	<0.001†
Black	392	(3.7)	784	(2.7)	
Asian	1,483	(14.0)	1,902	(6.4)	
Chinese	32	(0.3)	102	(0.4)	
Mixed	63	(0.6)	219	(0.7)	
Others	130	(1.2)	333	(1.1)	
Unknown	4,817	(45.6)	14,886	(50.5)	
Deprivation quintile, No. (%)					
(least deprived) 1	1,509	(14.3)	5,229	(17.7)	<0.001†
2	1,575	(14.9)	4,985	(16.9)	
3	2,177	(20.6)	6,149	(20.8)	
4	2,463	(23.3)	6,319	(21.4)	
(most deprived) 5	2,232	(21.1)	5,020	(17.0)	
Unknown	605	(5.7)	1,829	(6.2)	
CVD at index date, No. (%)	190	(1.8)	274	(0.9)	<0.001†
HF at index date, No. (%)	29	(0.30)	21	(0.07)	<0.001†
AF at index date, No. (%)	17	(0.16)	21	(0.07)	0.010†
Hypertension at index date, No. (%)	1,112	(10.5)	548	(1.9)	<0.001†
Prescription of NSAID, No. (%)	200	(1.9)	120	(0.4)	<0.001†
BMI (kg/m^2), No. (%)					
<20	332	(3.1)	2,191	(7.4)	<0.001††
20-24.9	1,542	(14.6)	9,452	(32.0)	
25-29.9	2,420	(22.9)	7,028	(23.8)	
30-34.9	2,104	(19.9)	2,868	(9.7)	
35-39.9	1,585	(15.0)	1,071	(3.6)	
≥40	1,575	(15.0)	644	(2.2)	
Unknown	1,003	(9.5)	6,277	(21.2)	
BMI (kg/m^2), median [IQR]	31	[26, 37.1]	24.9	[22.1, 28.7]	<0.001††
T2DM during follow-up, No. (%)	1,262	(11.9)	124	(0.4)	<0.001†
CKD during follow-up, No. (%)	182	(1.7)	126	(0.4)	<0.001†
Country, No. (%)					
England	8,516	(80.6)	23,661	(80.0)	0.688†
Northern Ireland	270	(2.6)	791	(2.7)	
Scotland	804	(7.6)	2,321	(7.9)	
Wales	971	(9.2)	2,758	(9.3)	

Abbreviations: CVD, cardiovascular disease; HF, heart failure; AF, atrial fibrillation; NSAIDS, non-steroidal anti-inflammatory drugs BMI, body mass index; IQR, interquartile range. Values are numbers (percentages) unless otherwise specified. P-values are from X^2 test (†) for proportions or t-test (††) for continuous variables (age, BMI) and Mann-Whitney U test (†††) for follow-up.

CHAPTER SEVEN: RESULTS
PREDICTORS OF CKD IN IGR

7.1 Introduction

This section describes the future CKD risk in patients following a diagnosis of IGR. The first part of the analysis examines the effect of the individual prognostic factors on the risk of CKD in a set of univariate analyses. Non-linear association between the continuous variables (age and BMI) and CKD were checked by fitting fractional polynomials (204). Survival probability (remaining free of CKD) was estimated and described using Kaplan-Meier survival estimates. A global test of the proportional hazards assumption (required for the Cox multivariable model) was undertaken for all covariates of interest, and additional log-log plots were used to check this assumption for specific covariates and to assess how best to enter BMI in the final model. A multivariable Cox regression was used to estimate the risk of CKD following a diagnosis of impaired glucose regulation (IGR). The goodness of fit of the final model was estimated by using the R-squared (R^2) statistic for time-to-event models developed by Royston and Sauerbrei (205).

7.2 Predictors of CKD in IGR - Study cohort characteristics

The study population included in this analysis consists of 10,561 patients with IGR and without CKD at baseline, followed-up for a maximum of 15 years. The median age of patients with CKD following a diagnosis of IGR was 36 years [interquartile range (IQR): 32 – 38 years] with 52.8% being female. The median follow-up was 1.0 year [IQR: 0.2, 2.7]. Follow-up for CKD patients was shorter because they were censored on the date patients developed CKD. Of the 10,561 patients with IGR, 182 (1.7%) developed CKD. During follow-up 41 (22.5%) patients with IGR who developed CKD progressed to type 2 diabetes (T2DM) (and were censored at this point).

The majority of patients with CKD following a diagnosis IGR were White (40.7%) followed by South Asian (15.4%). Information on ethnicity was missing for 39% of the IGR cohort who developed CKD. Missing data on ethnicity was categorised and a separate 'Unknown' category created. The most common body mass index (BMI) groups were 25 – 29.9 kgm^2 (26.4%) and 30 – 34.9 kg/m^2 (26.4%) and 3.9% of the IGR cohort who developed CKD had missing BMI data. Baseline characteristics of the IGR cohort split by those who did and did not have CKD are shown in Table 14.

Table 14: Demographic and clinical characteristics of the IGR cohort split by those who did and did not have CKD

	CKD		No CKD	
IGR (N=10,561)				
Age, median [IQR]	36	(32,38)	35	(31, 38)
Female, No. (%)	96	(52.8)	5,563	(53.6)
Years of follow-up, median [IQR]	1.0	(0.2, 2.7)	1.7	(0.8, 3.2)
Ethnicity, No. (%)				
White	74	(40.7)	3,570	(34.4)
Black	7	(3.9)	385	(3.7)
South Asian	28	(15.4)	1,455	(14.0)
Chinese	1	(0.6)	31	(0.3)
Mixed	1	(0.6)	62	(0.6)
Others	0	(0.0)	130	(1.3)
Unknown	71	(39.0)	4,746	(45.7)
Deprivation quintile, No. (%)				
(least deprived) 1	26	(14.3)	1,483	(14.3)
2	26	(14.3)	1,549	(14.3)
3	40	(22.0)	2,137	(20.6)
4	39	(21.4)	2,424	(23.4)
(most deprived) 5	38	(20.9)	2,194	(21.1)
Unknown	13	(7.1)	562	(5.7)
CVD at index date, No. (%)	3	(1.7)	187	(1.8)
HF at index date, No. (%)	0	(0.0)	29	(0.3)
AF at index date, No. (%)	1	(0.6)	16	(0.2)
Hypertension at index date, No. (%)	46	(27.3)	1,066	(10.3)
Prescription of NSAID, No. (%)	38	(20.9)	2,435	(23.5)
BMI (kg/m^2), No. (%)				
<20	1	(0.6)	331	(3.2)
20-24.9	21	(11.5)	1,521	(14.7)
25-29.9	48	(26.4)	2,272	(22.9)
30-34.9	48	(26.4)	2,056	(19.8)
35-39.9	27	(14.8)	1,558	(15.0)
≥40	30	(16.5)	1,545	(14.9)
Unknown	7	(3.9)	996	(9.6)
BMI (kg/m^2), median [IQR]	32	(26.7, 38.2)	31	(26, 37.1)
T2DM during follow-up, No. (%)	41	(22.5)	1,221	(11.8)
Country, No. (%)				
England	167	(91.8)	8,349	(80.4)
Northern Ireland	0	(0.0)	270	(2.6)
Scotland	6	(3.3)	778	(7.7)
Wales	9	(5.0)	962	(9.3)

Abbreviations: CVD, cardiovascular disease; HF, heart failure; AF, atrial fibrillation; NSAIDS, non-steroidal anti-inflammatory drugs; BMI, body mass index; IQR, interquartile range. Count of T2DM and CKD outcomes were for a maximum of 15 years following IGR diagnosis.

7.3 Risk of CKD in patients with IGR for risk factors of interest (univariate analysis)

The results of the univariate analyses are shown in Table 15. There was an increased risk of CKD with each additional year of age at IGR diagnosis (HR, 1.08; 95% CI, 1.04 to 1.12, p=<0.001). By comparison with the White ethnic group, patients with missing ethnicity were the only group showing a statistically significant difference in risk of CKD (HR, 0.68; 95% CI, 0.49

to 0.94, p=0.020). When compared with the most frequent BMI group (25 – 29.9), the only group to a statistically significant change in risk were patients with a BMI<20 (HR, 0.13; 95% CI, 0.02 to 0.97, p=0.046) and those with missing BMI (HR, 0.32; 95% CI, 0.14 to 0.71, p=0.005). Of the comorbidities only hypertension showed a statistically significant association (HR, 3.25; 95% CI, 2.32 to 4.54, p=<0.001). There was no statistically significant association between risk of CKD and gender, deprivation quintile, any known ethnic group, atrial fibrillation (AF), cardiovascular disease (CVD), heart failure (HF) or prescription of a non-steroidal anti-inflammatory drug (NSAID).

A pragmatic decision was made to combine groups with smaller numbers of IGR cases/outcomes for the final model to ensure that reliable estimates of risk would be obtained. Body mass index of <20 and 20 - 24.9 were combined to create a single category of <25 because of the relatively small numbers of cases in patients with BMI <20. The ethnic groups Chinese and Other were also combined because no CKD outcomes were observed in the 'Other' group.

Table 15: Risk of CKD in IGR – Univariate analysis

Factor	Categories	Number at risk	CKD outcomes	Hazard Ratio (HR)	95% Confidence Interval (CI)	*P*-Value
Age	(per 1 year)	10,561	182	1.08	(1.04, 1.12)	<0.001
Gender	Male (Reference)	4,902	86	1.00		
	Female	5,659	96	0.94	(0.70, 1.26)	0.692
BMI (kgs/m^2)	<20	332	1	0.13	(0.02, 0.97)	0.046
	20-24.9	1,542	21	0.67	(0.40, 1.12)	0.123
	25-29.9 (Reference)	2,420	48	1.00		
	30-34.9	2,104	48	1.17	(0.79, 1.75)	0.433
	35-39.9	1,585	27	0.86	(0.54, 1.38)	0.533
	≥40	1,575	30	0.96	(0.61, 1.52)	0.872
	Missing	1,003	7	0.32	(0.14, 0.71)	0.005
Townsend Quintile	1 (Reference) (least deprived)	1,509	26	1.00		
	2	1,575	26	0.99	(0.58, 1.71)	0.973
	3	2,177	40	1.10	(0.67, 1.80)	0.713
	4	2,463	39	0.98	(0.60, 1.61)	0.936
	5 (most deprived)	2,232	38	1.06	(0.64, 1.74)	0.829
	Missing	605	13	1.36	(0.70, 2.64)	0.367
Ethnicity	White (Reference)	3,644	74	1.00		
	South Asian	1,483	28	1.06	(0.69, 1.64)	0.781
	Black	392	7	1.01	(0.46, 2.19)	0.983
	Chinese	32	1	1.98	(0.28, 14.25)	0.498
	Mixed	63	1	0.90	(0.13, 6.47)	0.916
	Other	130	0	N/A		
	Missing	4,817	71	0.68	(0.49, 0.94)	0.02
CVD	CVD	190	3	0.96	(0.31, 3.00)	0.942
Heart Failure	HF	29	0	N/A		
Atrial Fibrillation	AF	17	1	3.55	(0.50, 25.31)	0.207
Hypertension	Hypertension	1,122	46	3.25	(2.32, 4.54)	<0.001
NSAID	NSAID	2,473	38	0.94	(0.66, 1.34)	0.736

Abbreviations: CKD, chronic kidney disease; IGR, impaired glucose regulation; HF, heart failure; CI, confidence interval

7.4 Checking for non-linear relationships between the continuous variables and the outcome

A non-linear association between the two continuous covariates (age and BMI) and risk of CKD was checked for using an automated method (mfp in STATA) (197) for fitting fractional polynomials developed by Royston and colleagues (194). A linear association was found to be the best fit for BMI and age.

7.5 CKD-free survival following a diagnosis of IGR

Table 16, describes the probability of remaining free of CKD up to 15 years following IGR diagnosis. A decision was made to present the 5 year survival probability (Table 17) of CKD following IGR diagnosis as fewer than 10% of the original cohort was still under observation after the fifth year. Rich and colleagues (198) point out that Kaplan-Meier (K-M) estimates are most accurate when most patients are still present in the study. Furthermore, K-M estimates can be misleading and should be interpreted with caution if only a subset of the population is included. During the first year, 7269 (69%) patients were still being followed-up, in the following year this was reduced to 4487 (43%), at five years 1098 (10%) patients remained and 1 (0.009%) at 15 years. The risk of CKD by 2 years following IGR was 2% (0.98; 95% CI, 0.98 to 0.99). At 5 years follow-up the risk of CKD was approximately 4% (0.96; 95% CI, 0.96 to 0.97). The K-M estimate shows that if IGR persists over a sustained period, the higher the risk of CKD.

Table 16: Fifteen years CKD risk estimates in patients with IGR – survival probability

Following diagnosis of IGR (Year)	Population at risk	CKD events	Survival probability	95% (CI)	
Total	10561				
1	7269	95	0.99	0.99	0.99
2	4487	33	0.98	0.98	0.99
3	2834	12	0.98	0.98	0.98
4	1760	23	0.97	0.97	0.98
5	1098	9	0.96	0.96	0.97
6	664	7	0.96	0.95	0.96
7	438	3	0.95	0.94	0.96
8	270	0			
9	156	0			
10	77	0			
11	45	0			
12	20	0			
13	7	0			
14	2	0			
15	1	0			

Abbreviations: IGR, impaired glucose regulation; CKD, chronic kidney disease; CI, confidence interval

Table 17: Five years CKD risk estimates in patients with IGR – survival probability

Following diagnosis of IGR (Year)	Population at risk	CKD events	Survival probability	95% (CI)	
Total	10561				
1	7269	95	0.99	0.99	0.99
2	4487	33	0.98	0.98	0.99
3	2834	12	0.98	0.98	0.98
4	1760	23	0.97	0.97	0.98
5	1098	9	0.96	0.96	0.97

Abbreviations: IGR, impaired glucose regulation; CKD, chronic kidney disease; CI, confidence interval

7.6 Testing proportional hazards assumptions – Five-year follow-up

The proportional hazards assumption for each predictor along with the overall test for the risk of CKD (Table 18) was checked using the Schoenfeld residuals (estat phtest in STATA) (200). Proportionality was also checked by plotting the log-log survival curves for selected covariates (

14 - 16). The overall test result indicates that the hazard proportional assumption was met (Chi-square [χ^2] = 17.04, p= 0.650). However, the log-log plot for BMI did not show a consistent

relationship between increasing BMI and CKD risk, so BMI was entered into the analysis as a set of distinct BMI groups.

Table 18: Proportional hazard assumption test for each predictor for the risk of CKD in IGR – Five year follow-up

Predictors		Chi^2	*P*-Value
Age		0.00	0.995
Gender			
Male	Reference		
Female		1.66	0.198
BMI (kgs/m^2)			
<25		0.47	0.494
25-29.9	Reference		
30-34.9		0.82	0.365
35-39.9		0.34	0.557
≥40		0.95	0.330
Missing		0.19	0.665
Deprivation quintile			
Least deprived	Reference		
2		0.02	0.884
3		0.24	0.621
4		0.52	0.469
Most deprived		4.00	0.046
Missing		0.13	0.723
Ethnicity			
White	Reference		
South Asian		0.54	0.463
Black		4.71	0.030
Chinese/Other		0.16	0.687
Mixed		4.26	0.039
Missing		1.21	0.272
CVD		0.05	0.819
HF		N/A	
AF		N/A	
Hypertension		0.68	0.408
NSAID		0.02	0.901
Overall test		17.04	0.650

Abbreviations: BMI, body mass index; CVD, cardiovascular disease; HF, heart failure; AF, atrial fibrillation; Chi^2, Chi-Square test
Note: The covariates HF and AF do not have χ^2 and p-values reported because no failures were observed in the first 5 years.

7.7 Risk of CKD in patients with IGR for risk factors of interest (multivariate analysis)

The results of the Cox model are presented in Table 19. Heart failure and atrial fibrillation were removed from the analysis because no CKD outcomes were observed in the first 5 years. If they were left in, the effect of these two diseases would appear protective which is not clinically plausible.

There was a 7% increased risk of CKD for each additional year of age at IGR diagnosis (HR, 1.07; 95% CI, 1.03 to 1.11). Hypertension remained predictive of CKD in IGR (HR, 2.84; 95% CI, 1.99 to 4.05). There was no statistically significant association with BMI categories, ethnicity or gender. There was no statistically significant association between increasing deprivation or the two remaining predictors (CVD and NSAID) and CKD risk.

Table 19: Cox proportional CKD risk estimates in IGR patients - Multivariate analysis (five-year follow-up)

Factor	Categories	Number at risk	Number with CKD	Hazard Ratio (HR)	95% Confidence Interval (CI)	*P*-Value
Age	(per 1 year)	10,561	172	1.07	(1.03, 1.11)	0.001
Gender	Male (Reference)	4,902	80	1.00		
	Female	5,659	92	1.10	(0.80, 1.50)	0.551
BMI (kg/m^2)	<25	1,874	21	0.67	(0.40, 1.13)	0.138
	25-29.9 (Reference)	2,420	45	1.00		
	30-34.9	2,104	46	1.12	(0.74, 1.70)	0.590
	35-39.9	1,585	25	0.81	(0.49, 1.33)	0.397
	≥40	1,575	28	0.91	(0.56, 1.49)	0.705
	Missing	1,003	7	0.45	(0.20, 1.01)	0.054
Townsend Quintile	1 (Reference)	1,509	25	1.00		
	2	1,575	25	0.98	(0.56, 1.71)	0.951
	3	2,177	36	1.04	(0.62, 1.73)	0.895
	4	2,463	39	1.01	(0.61, 1.68)	0.973
	5	2,232	35	1.03	(0.61, 1.73)	0.913
	Missing	605	12	1.33	(0.66, 2.69)	0.423
Ethnicity	White (Reference)	3,644	70	1.00		
	South Asian	1,483	26	1.08	(0.68, 1.72)	0.748
	Black	392	7	0.96	(0.43, 2.11)	0.910
	Chinese/Other	162	1	0.35	(0.05, 2.52)	0.296
	Mixed	63	1	0.92	(0.13, 6.61)	0.931
	Missing	4,817	67	0.72	(0.51, 1.00)	0.052
CVD	CVD	190	3	0.80	(0.25, 2.52)	0.703
Hypertension	Hypertension	1,122	45	2.80	(1.96, 4.00)	<0.001
NSAID	NSAID	2,473	35	0.81	(0.55, 1.18)	0.266

Abbreviations: BMI, body mass index; CVD, cardiovascular disease; HF, heart failure; AF, atrial fibrillation; CI, confidence interval

7.8 Model checking: goodness of fit

The proportion of variation explained by the model (R^2) statistic was 0.24. This indicates that the final model accounted for 24% of the variation in the outcomes observed.

7.9 Conclusion

Patients with hypertension were approximately 3 times more at risk of CKD and for each additional year of age at IGR diagnosis, CKD risk increased by 7%. The overall proportion of variation explained by the model was 24% (R^2 0.24). The Kaplan-Meier (K-M) estimates suggest that 2% of IGR patients were diagnosed with CKD by 2 years increasing to 4% after 5 years follow-up.

CHAPTER EIGHT: DISCUSSION

8.1 Introduction

The primary focus of this thesis was to provide a robust estimate of the incidence of CKD stage 1-5 in young adults aged 18 to 40 years with IGR compared to those with normoglycaemia, and to estimate the risk of CKD 1 – 2 and 3 – 5 separately in the IGR cohort. Additionally, a model presenting risk factors associated with future CKD outcome was devised to understand the course and factors associated with future CKD development in young adults with diagnosed IGR. A systematic review was conducted to distil available evidence exploring the incidence of CKD in young adults with IGR compared to those with normoglycaemia, and investigate whether any increased risk occurs only after they develop T2DM. An analysis of routinely recorded clinical data derived from a large UK primary care database investigated the incidence of CKD in young adults with IGR. This chapter summarises the key findings of these analyses and discusses their implications.

8.2 Summary of key findings

The first objective of this thesis was to systematically identify and summarise the best available evidence showing whether the risk of CKD is elevated in young adults with IGT/IFG compared to normoglycaemia or T2DM. The systematic review shows a paucity of studies exploring the incidence of CKD in young adults with IGT/IFG. The review showed that existing evidence does not allow quantification of CKD risk in young adults aged 18 to 40 years with IGT/IFG compared to normoglycaemia or T2DM. Pooled estimates of CKD risk and a meta-analysis were not possible because most studies did not report separate results in this age group (Chapter 3).

The second objective investigated the incidence of CKD in a retrospective matched cohort analysis using the THIN database and found that young adults with IGR were four times more at risk of developing CKD than individuals with normoglycaemia. After adjusting for confounders, the effect of CKD risk was attenuated but was still 2.6 times higher in individuals with IGR than those with normoglycaemia (Chapter 6). Additionally, the incidence of CKD stage (3 – 5) was approximately 4 times higher than the incidence of CKD stage (1 – 2) following a diagnosis of IGR (Chapter 6).

The third research theme examined the incidence and prevalence of recorded IGR in young adults in the THIN database. The overall incidence of IGR found in this study was 72.2 cases per 100,000 person-years of follow-up. Incidence was higher in the age group 26 to 40 years than those aged 18 to 25 years and was higher in females than males. Incidence was higher in patients living in the most deprived areas than those in the least deprived areas and higher in Black and South Asian than White ethnic groups. Incidence increased annually and was approximately nine times higher in 2014 than 2000 (Chapter 6). Additionally, the prevalence of diagnosed IGR increased consistently from 2000 to 2014. Prevalence was consistently higher in females than males and was higher in the age group 26 to 40 years than those aged 18 to 25 years and increased consistently in both age groups and both genders over the same period (Chapter 6).

In the fourth and final research theme, future risk of CKD in patients with IGR was derived in a risk model. The aim of this model was to identify likely risk factors and determine their values in predicting incident CKD in young adults with IGR. Of all the modifiable risk factors included in the model only hypertension was significant. Patients with hypertension were approximately three times more at risk of CKD. Similarly, the effect of age was significant, there was a 7% increased risk of CKD for each additional year of age at IGR diagnosis. The Kaplan-Meier (K-

M) estimates show a 2% increased risk of CKD by 2 years following IGR, which increased to 4% after 5 years follow-up. The proportion of all the variation explained by the model was 24% (R^2 0.24). Although the amount of variation in the outcome observed is moderate, the model found that age and hypertension were significant risk factors in predicting incident CKD (Chapter 7).

8.3 Strengths and limitations

The present study contributes novel findings by providing reliable estimates of the incidence of CKD in young adults with IGR compared to those with normoglycaemia by systematically reviewing available literature and by analysing data from large generalisable cohort of primary care patients.

8.3.1 Systematic review

Systematic reviews are increasingly recognised as the most robust and reproducible approach utilised to identify, critically appraise and synthesise all available evidences in order to provide an empirically reliable answer to a defined research question (206). However, systematic reviews are laborious and time consuming which relies on the identification of all the relevant articles. Although the search method was rigorous and systematic, retrieval of all the relevant articles could not be assured. This review was conducted according to the PRISMA (207) and MOOSE (151) guidelines. Furthermore, this review followed a pre-defined peer-reviewed protocol published in an Open Access scientific journal (152). From a review perspective, the strengths were that a comprehensive search strategy which was not restricted to the English language, geographical area or length of follow-up and a broad range of markers were used to identify patients with CKD. Additionally, two reviewers independently screened titles and abstracts, identified eligible studies and assessed quality of the included studies. Furthermore,

only observational studies with incident CKD were included, because of their large sample size, high rate of follow-up and frequency of CKD, which are more likely to be representative of the population at risk. The main limitations were because of the restrictive age range, quantification of CKD risk was not possible because only one study was found providing CKD risk estimates in persons aged 18 to 40 years with IGR. Furthermore, it was not possible to conduct a meta-analysis because sufficient studies were not available, therefore a generalisable and precise estimates of CKD could not be presented. Additionally, results of this one study cannot be generalised because of the characteristics of the study population (Pima Indians).

8.3.2 Analysis of the THIN data

The strengths and weaknesses of utilising routinely collected electronic primary care data are discussed in (Chapter 2). In summary, the strengths include the large sample size of the database (THIN contains medical records of over 12 million patients) which is approximately 6% of the UK population (185). Generalisability was ensured by the fact that the dataset includes similar patients to those seen in primary care across the UK. This reflects the fact that data are collected in a non-interventional fashion during routine clinical practice. Internal validity (reducing the effect of confounders) was ensured by matching patient's age, sex and general practice. Although the study was retrospective, data were collected prospectively, hence eliminating recall bias. Ascertainment of exposure and outcome of interest was carried out by both Read coded diagnosis and clinical measurements (e.g. HbA1c, ACR lab results). Furthermore, a single eGFR was used to identify patients with CKD in addition to patients with diagnosed CKD with a repeat serum creatinine as defined by NICE guidelines (9). This was undertaken because it is likely that patients suspected of being at high risk of CKD (e.g., hypertension, raised BMI) would have their serum creatinine recorded and eGFR calculated on one occasion and not necessarily have a

repeat serum creatinine test. Additionally, the completeness of lab results has significantly improved since it is now directly fed into GP databases through Path Link populating patient records. Therefore, improving ascertainment of IGR and CKD for this study and providing a reliable estimate of incident CKD in young adults with IGR. The main limitations relate to the identification of IGR, which depends on GPs doing blood tests. Therefore, the incidence and prevalence of IGR is an underestimate of the true incidence and prevalence and it is also biased as GPs are more likely to do blood tests for glucose in patients in whom they suspect diabetes, (e.g. high BMI, older age, some ethnic groups). This will exaggerate the association between IGR and these risk factors. Similarly, incidence and prevalence of recorded CKD also depends on GPs doing blood tests. This may also underestimate the frequency of CKD and may also be biased because GPs are more likely to do blood tests for CKD in patients in whom they suspect CKD. This includes those with diabetes, hypertension, certain ethnic groups and those with IGR. This may also increase the relationship between IGR and CKD. For this study both laboratory evidence of reduced eGFR and Read code were used to identify incident CKD patients with only Read coded CKD or only laboratory evidence of CKD would not be misclassified.

There is a lack of data on some important confounding variables (e.g., ethnicity, BMI) pre Quality and Outcomes Framework (QOF) [prior to 2004]. Additionally, before lab-linkage practices lab results were entered manually and in some cases only abnormal results may have been entered into the GP system, this could underestimate the incidence of CKD. Furthermore, since 2006, general practitioners (GPs) in England have been incentivised to keep a register of patients with CKD stage 3-5 through QOF. This may have increased diagnosis entered in THIN in the later years (post 2006).

There is also potential for ascertainment bias which may arise when patients who have frequent laboratory tests may be more likely to have IGR diagnosed and also may be more likely to have CKD diagnosed. It can be seen that the prevalence of diagnosed IGR increased 7 fold, from 2000 to 2014 (Table 8). If more laboratory tests were being undertaken we would expect that the number of patients diagnosed with IGR would also increase. This seems to have happened, with a 15 fold increase in diagnosed IGR from 2000 to 2014 (Table 8). Additionally, the number of IGR cases detected from laboratory tests increases 20 fold over the study period (Table 13c). This is likely to be due to increased laboratory testing and increased recording of laboratory tests in electronic patient records. Increased laboratory testing and increased recording of laboratory tests could also lead to an increase in diagnosis of CKD. Moreover, before when GP was receiving lab results abnormal tests such as glucose on a pre-diabetic range may have been overlooked and not entered onto the database. Additionally, the median follow-up of this study was approximately two years. This may have been insufficient to allow for the development of CKD among those with IGR. Furthermore, the declining number of CKD diagnoses observed in the later years of the study may be attributed largely to IGR patients being censored (de-registration, death and patients progressing to diabetes). Therefore, incidence of CKD cannot be reliably estimated in the later years of this study. Lastly, application of the predictors of CKD in IGR model developed as part of this thesis may not be appropriate because the relationship between predictors and CKD may change from the years the data was derived to develop this model (January 2000 to February 2015). This may be caused by changes in the recording of IGR and CKD or by changes in diagnosis or treatment strategies for CKD or IGR. Changes in the completeness of data on ethnicity and BMI may also change the relationship between predictors and CKD

8.4 Demographic and clinical characteristics of the IGR cohort

The characteristics of the study cohort and their clinical estimates at the point of IGR diagnosis are presented in Table 13 (Chapter 6). The median age of the study cohort was 35 years and 53% were female. Patients with IGR were more likely to be of White ethnicity followed by South Asians. Additionally, patients with IGR were more likely to be from areas of greater deprivation and were more likely to have a BMI of 25 to 29 kg/m^2. Atrial fibrillation, heart failure and cardiovascular disease were found in 0.16%, 0.3% and 1.8% of the IGR population respectively. Furthermore, hypertension was found in more than 10% of the IGR cohort and 1.9% were on non-steroid anti-inflammatory drugs.

8.5 Clinical significance of the findings presented in this thesis

Diabetes is related to a wide range of macrovascular (coronary artery disease, peripheral vascular disease and stroke) and microvascular (retinopathy, neuropathy and nephropathy) complications (208). Kidney disease is frequently diagnosed during routine consultation and generally occurs in people with diabetes. The United Kingdom Prospective Diabetes Study (UKPDS trial) demonstrated a high incidence (29%) of renal impairment (eGFR<60 ml/min per 1.73 m^2 or a doubling of serum creatinine) after patients developed T2DM (209). Patients identified with IGR are also at increased risk of progressing to overt diabetes (210). Results of this study indicate that the trend in the prevalence and incidence of IGR is set to continue to increase. The trend shows a consistent increase in prevalence and incidence throughout the calendar period (2000-2014) in both males and females which are particularly alarming because this may lead to an increased number of adults aged 18 to 40 progressing to diabetes and its complications.

Furthermore, there is also evidence suggesting that elevated blood glucose associated with IGR significantly increases the risk of kidney abnormalities resulting in CKD (5). The current study appears to confirm that elevated blood glucose associated with IGR increases the risk of CKD stage (3 – 5) by approximately 4 times than the risk of CKD stage (1 – 2). Furthermore, the current study also confirms result of the Framingham study. The unadjusted incidence of CKD in patients with IGR was approximately four times higher than patients with normoglycaemia. After adjustment for potential confounders, risks was slightly attenuated, the incidence was reduced to approximately three times higher. The annual incidence of CKD in patients with IGR may be low (e.g., 1 in 1000 person-years), if treatment was to halve the incidence from 1 to 0.5 per 1000 person years it means that 2000 patients must be treated per year for one extra patient to benefit. This may not change the approach to care by healthcare professionals engaged in the management of these patients but should at least provide some perspective that raised blood glucose over a sustained period of time may cause microvascular complications and may warrant the need for patients with IGR to have their kidney function checked routinely.

The results presented in Tables 15 and 19 of the CKD predictors in IGR model development (Chapter 7) shows that hypertension was the only comorbidity to show a statistically significant association in univariate analysis which remained predictive in multivariate analysis. Patients with hypertension at IGR diagnosis were approximately four times more at risk of CKD in the univariate analysis. Risk of CKD was slightly reduced to approximately three times in multivariate analysis. Hypertension has been shown to be a strong independent modifiable risk factor associated with an increased risk of kidney disease. This was demonstrated in a retrospective cohort study of 118,924 Taiwanese participants without diabetes and CKD at baseline. The study reported a 23% increased risk of CKD after adjustment in participants with

hypertension (6). Considering the results of the study and existing evidence, hypertension remains one of the most potent risk factor for CKD development but also easily treatable if adequately managed in high risk groups. Good control of hypertension remains one the most efficient way in managing CKD.

8.6 Comparison with existing literature

8.6.1 Systematic review

To the best of our knowledge, no systematic review evaluated the risk of CKD in young adults aged 18 to 40 years with IGT/IFG compared to normoglycaemia or T2DM. According to a recently published systematic review, patients with prediabetes are at increased risk of developing CKD. Echouffo-Tcheugui and colleagues (184) examined nine cohort studies to estimate the incidence of CKD in adults aged (≥ 18 years old) with prediabetes, defined as (IGT/IFG). The pooled summary estimate in a random effect meta-analysis shows that the risk of CKD after adjustment for established CKD risk factors was 11% higher in individuals with prediabetes than those with normoglycaemia. Additionally, a systematic review examined the effect of metabolic syndrome and the risk of developing CKD, proteinuria or micro-albuminuria. This found that as an individual component of metabolic syndrome, the pooled summary estimate for impaired fasting glucose (IFG) shows a 16% increased risk of CKD (211). In contrast, several studies (166, 212) examined the effect of IGR and risk of CKD where IGR was included as a component of the metabolic syndrome. A prospective study conducted by Tozawa and colleagues (166) to examine the effect of metabolic syndrome on the incidence of CKD determined by (eGFR<60ml/min/1.73m^2) in Japanese adults aged (≥19 years), found that individuals with IGR had a 22% increased risk of developing CKD after adjusting for age, sex, current cigarette smoking and alcohol drinking habits. Additionally, Kurella and colleagues

(212) reported an 11% increased risk of CKD in individuals with IFG. In contrast, Nelson and colleagues (167) in a population of 490 subjects (183 male, 307 female), reported a weak association between IGR and development of CKD (OR: 0.8, 95% CI; 0.4 to 1.5). Additionally, Nelson and colleagues (102) in a subsequent study followed 194 Pima Indians at intervals of 6 to 12 months for 4 years to analyse the changes in glomerular filtration rate that occur during progression of renal disease. Six groups were recruited with predefined characteristics: 31 had normal glucose tolerance; 29 with impaired glucose tolerance and 30 newly diagnosed diabetes. During the 4 year follow-up the glomerular filtration rate increased by 14% (p=0.008) in subjects with IGR. This was confirmed in another study; GFR was serially measured for 48 months in 26 Pima Indians with IGR and 27 with normal glucose tolerance. At baseline, the mean GFR was normal (<90ml/min/1.73m^2). The baseline GFR increased by 20% in IGR compared to 14% in NGT, showing an association with increased GFR and elevated glucose tolerance (163). These studies, however, did not report separate results in the age group 18 to 40 years.

8.6.2 Incidence of IGR in young adults

Several studies have used large generalisable sample of UK general practices data. However, to the best of our knowledge, no previous study using the THIN database has explored the incidence of diagnosed IGR in young adults aged 18 to 40 years. A population based cohort study examining the incidence rate of diabetes and prediabetes in a South Asian cohort aged (≥20 years) from Chennai was carried out by Anjana and colleagues (213). The researchers used a sub-set of the Rural Epidemiology Study (CURES) cohort study. Prediabetes was defined as fasting plasma glucose of (5.6 to 6.9 mmol/L) or an oral glucose tolerance test of (7.8 to 11.0

mmol/L). After a median follow-up of 9.1 years, the incidence rate of prediabetes was 7.1 cases per 1000 person-years of follow-up. The incidence rate reported in this study is lower than those reported in the Biracial Cohort (POP-ABC) study, the incidence of prediabetes was 12.5 cases per 100 person-years of follow-up in adults aged (≥18 years) (214). Conversely, in a study using data from the General Practice Research Database (GPRD), the annual incidence rate of IGT/IFG, defined by a Read code suggesting IGT/IFG and reported measurements of blood and urine glucose, in adults aged (≥20 years) increased from 17 cases per 100,000 person-years in 2000 to 31 per 100,000 person-years in 2004 (78). The incidence rate of IGR in 2004 presented in this study is higher (44.9 cases per 100,000 person-years) than the incidence rate reported by Gillett and colleagues in 2004 (31 cases per 100,000 person-years). The higher incidence rate observed in this study may be associated with a broader diagnostic criteria used to define IGR (Chapter 4 and Appendix 11). Incidence rates should be interpreted with caution because these studies were conducted in specific populations and separate results in the age group (18 to 40 years) were not provided. Furthermore, incidence rates may be affected by population structure (e.g. population with a higher proportion of older people with IGR may have a higher incidence of CKD than a younger population with IGR).

8.6.3 Prevalence of diagnosed IGR in young adults

Prevalence estimates especially differ from other published work mainly because not all patients are tested for IGR. The prevalence in the present study is lower than the prevalence reported in a report published by the National Cardiovascular Intelligence Network (NCVIN) (77). The NCVIN analysis used a combined dataset (2003 to 2013) from the Health Survey for England (HSE) to determine the prevalence of non-diabetic hyperglycaemia also known as pre-diabetes or impaired glucose regulation in adults aged (≥16 years). Non-diabetic hyperglycaemia was

defined as an HbA1c value between 6.0 – 6.4% (42 – 47 mmol/mol) excluding those diagnosed with diabetes. The prevalence was 2.6% in the 16 to 39 years age group. In contrast, Mainous and colleagues (76) estimated prevalence of pre-diabetes in adults aged (≥16 years) using data from HSE for the years 2003, 2006, 2009 and 2011. Pre-diabetes was determined by the ADA (68) definition, a glycated haemoglobin (HbA1c) value of 5.7 – 6.4% (39 – 47 mmol/mol). They reported a pre-diabetes prevalence of 11.6% in 2003 rising to 35.3% in 2011. Prevalence of pre-diabetes in the age group 16 to 30 years was 5% in 2011. Prevalence of pre-diabetes reported by Mainous and colleagues was higher than the prevalence reported in the present study. The difference in prevalence may be explained by the definitions used to determine IGR. Mainous and colleagues used a lower HbA1c 5.7 – 6.4% (39 – 47 mmol/mol) range, whereas this study used the NICE (65) recommended HbA1c cut-off value of 6.0 – 6.4% (42 – 47 mmol/mol). Additionally, data from Project STAND (Sedentary Time and Diabetes), a randomised controlled trial of obese young adults aged 18 to 40 years recruited within Leicestershire and South East Midlands, the prevalence of impaired glucose metabolism was 18.1% (75). In the Australian Diabetes, Obesity and lifestyle study (AusDiab), a cross-sectional survey across six states and the Northern Territory, the prevalence of IGR was 16.4% in adults aged (≥25 years) (215). Furthermore, the sole use of HbA1c as diagnostic criteria for IGR may mean that some patients with diabetes are included and therefore overestimate frequency of IGR. Incorrectly classifying patients with diabetes as IGR may increase the association with risk of CKD. However, incorrectly classifying those with normal glucose regulation as IGR may dilute the association between IGR and risk of CKD in this population. On the other hand, the STAND project used both an oral glucose tolerance test (OGTT) and HbA1c (6.0 – 6.4%) measurements to ascertain impaired glucose metabolism. Additionally, the AusDiab survey used only an OGTT to determine IGT/IFG. For the purpose of this thesis a broader diagnostic criteria (Chapter 4)

was used to determine IGR. Furthermore, the present study population was larger (approximately 3.5 million) compared to NCVIN (54,644), Mainous and colleagues (20,139) STAND (193) and AusDiab (11,247).

8.6.4 Incidence of CKD in young adults with IGR compared to those with normoglycaemia

To the best of our knowledge, this study was the first population based study utilising data from a large cohort of general practices and their patients from across the UK, and contemporary data, to investigate the incidence of CKD in young adults aged 18 to 40 years with IGR compared to those with normoglycaemia. Several studies have been conducted internationally exploring whether IGR as an individual component of a cluster of metabolic abnormalities is associated with an increased risk of CKD development determined by an (eGFR<60ml/min/1.73m2) or an albumin creatinine ratio (ACR) (6, 166, 170). Similarly, Watanabe and colleagues (7) in the Niigata Preventive Medicine Study, examining the risk of CKD in individual components of metabolic syndrome, found the incidence of IGR was approximately twice as high in subjects with IGR as those with normoglycaemia. Meigs and colleagues (162) in the Framingham Offspring study provided separate results for males and females, the odds of CKD development was 6% and 7% higher in males and females respectively with IGR compared to normoglycaemia. These studies, however, did not report separate results in persons aged 18 to 40 years with IGR and risk of CKD, but rather results were presented for the population as a whole. Therefore, quantification of CKD risk was not possible in the age group 18 to 40 years. Nevertheless, the results presented in this thesis appear to be consistent with previous studies showing an elevated risk of CKD in adults with IGR compared to those with normoglycaemia.

8.6.5 Incidence of CKD by each category (stages 1 – 2/3 – 5) in the IGR cohort

To the best of our knowledge this is the first population based study examining the burden of CKD by each category (stages 1 – 2/3 – 5) in young adults aged 18 to 40 years with IGR. In a community based 10 – year follow-up study, using data from an annual health check in Japan, Yamagata and colleagues (216) examined the risk factors associated with incident CKD in adults aged ≥40 years without evidence of CKD stage (1 – 2/3 – 5) at baseline. During the follow-up period, the hazard of developing CKD stage (1 – 2) in individuals with IGR increased by 21% in men (HR: 1.21; CI, 1.08 to 1.35) and 19% in women (HR: 1.19; CI, 1.05 to 1.35). In contrast, there was not a statistically significant association between CKD stage (3 – 5) in both males and females with IGR, the hazard were (HR: 0.86; CI, 0.79 to 0.93) and (HR: 0.80; CI, 0.76 to 0.85) respectively. This study, however did not report separate results in the age group 18 to 40 years. The follow-up period for this study was 10 years compared a shorter length of follow-up for individual patients in the current study (median 2 years approximately). In their study, Yamagata and colleagues included 113, 764 individuals aged (≥40 years) and 33% were male compared to a study population of over 40,000 individuals aged 18 to 40 years and 46.4% male in the current study.

8.6.6 Predicting future CKD outcome in young adults with IGR

Electronic patient records of patients registered with a general practice and contributing to the THIN database was used to develop a prediction model to quantify the effect of risk factors in determining incident CKD in young adults aged 18 to 40 years with diagnosed IGR. The K-M model was used to describe the risk of CKD over time following IGR diagnosis. The present prediction model used a broad range of diagnostic criteria (Read code and clinical measurements) to identify cases, outcomes and risk factors. In a recently published prediction

model, the incidence of CKD in patients with type 2 diabetes and free from CKD at baseline were determined using data derived from the Ongoing Telmisartan Alone and Ramipril Global Endpoint (ONTARGET) clinical trials (217). Two prediction models were developed (laboratory and clinical models). The clinical model was compared to the present study because this model contained comparable risk factors. This study used only urinary albumin creatinine ratio and eGFR (<15 ml/min per 1.73 m^2) to ascertain outcome compared to a range of clinical measurements used in the present study to define the outcome of interest. Risk factors included in this study differs from the present study. Some important predictors in this study (i.e. BMI and ethnicity) had limited detailed group which could have markedly different risks. The overall performance of the model (R^2=10.73%) reported in this study is lower compared to the present study (24%). It should be noted that this study included individuals aged (≥55 years), therefore like for like comparison cannot be made. Furthermore, predictors of CKD risk in the present study are likely to differ from the above study. The reason for this disparity may be attributed to the inclusion of adults aged 18 to 40 years in the present study compared to an older age group in the ONTARGET study. Furthermore, Riphagen and colleagues (218) developed a 10 year risk prediction model to determine the risk of microalbuminuria and increase in serum creatinine in patients with T2DM using data from the Zwolle Outpatient Diabetes Project Integrating Available Care (ZODIAC) study. Some risk factors included in this study are similar to those included in the current study. However, the statistical significance of each risk factor was not tested in univariate analyses before inclusion in the final model. The Cox model shows a 52% increased risk of microalbuminuria and a 45% increased risk of renal function loss determined by a 50% increase in baseline serum creatinine for each 10 additional years of age at T2DM diagnosis. In the current study, there was a 7% increased risk of CKD for each additional year of age at IGR diagnosis. In this study males showed a 63% increased risk of microalbuminuria

compared to females. However, there was no statistically significant association between gender and an increase in serum creatinine. BMI was presented as a continuous variable, there was a 7% increased risk of renal function loss for every unit increase in BMI. For each unit increase in systolic blood pressure the risk of microalbuminuria and renal function loss was 16% and 12% respectively. Hypertension in the current study showed a statistically significant association in the univariate analysis and remained predictive of CKD in IGR (HR, 2.84) in the multivariate analysis. Macrovascular diseases were tested as a single group, both risks of microalbuminuria (HR: 1.52) and renal function loss (HR: 1.22) were significant. There was no statistically significant association between risk of CKD and atrial fibrillation, cardiovascular disease, heart failure or prescription of a non-steroidal anti-inflammatory drug in the current study. The study contained data for 1,143 patients with a mean age 68(+/-12) and 42.8% were male compared to a study population of 10,561 patients with IGR and 46.4% male in the current study.

8.6.7 Ascertainment of exposure and outcome compared to other studies

Other studies have investigated the risk of CKD in people with impaired glucose regulation (5, 169). However, the method used to ascertain exposure and outcome of interest vary widely between studies. The current study used a wide range of clinical measurements obtained from laboratory results and Read coded diagnosis to identify exposure/ outcome of interest and relevant covariates compared to other studies (Chapter 4). In a prospective study, Fox and colleagues (5), investigated the effect of worsening glucose in non-diabetic patients and risk of CKD, used only a 2 hour oral glucose tolerance (OGTT) test to identify patients with IGR, kidney function on the other hand was estimated by an (eGFR<60 ml/min per 1.73 m^2), from serum creatinine results and calculated using the modification of diet in renal disease (MDRD) equation. Additionally, Kurella and colleagues (212) using data from a longitudinal prospective

study of cardiovascular disease risk in non-diabetic adults (Atherosclerosis Risk In Communities [ARIC] study) examined the effect of metabolic syndrome and risk of CKD development. Metabolic traits were defined by the National Cholesterol Education Program (NCEP) guidelines. Patients with impaired fasting glucose were identified by fasting glucose (≥110mg/dL) alone and the outcome of interest, CKD was defined by an (eGFR<60 ml/min per 1.73 m^2) using serum creatinine and calculated by the MDRD equation.

8.7 Future research

A diagnosis of IGR in young adults aged 18 to 40 years may be an incidental discovery when screening for comorbidities in primary care. This study has shown that young adults with IGR are at increased risk of developing CKD. There are a number of areas that could be investigated in future research:

- A randomised controlled study to investigate whether hypertension treatment can prevent the development of kidney disease among those with IGR aged 18 to 40 years

- Qualitative research with GPs might help understand biases in ascertainment and recording of IGR and CKD by understanding what risk factors might cause GPs to do blood tests for IGR or CKD in young adults.

- A clinical trial investigating whether glucose lowering drugs versus glycaemic control prevent the development of CKD following IGR diagnosis

- Follow up study of young adults with IGR to determine incidence of CKD

- Further research to determine whether lifestyle interventions strategy could potentially prevent or delay progression of IGR in young adults aged 18 to 40 years.

- A cohort study to validate the sole use of HbA1c (6.0 – 6.4%) against a combination of HbA1c and OGTT as a method to reliably diagnose IGR in young adults aged 18 to 40 years.
- A cohort study using the THIN database to identify which IGR patients following introduction of the NHS Diabetes Prevention Programme are at highest risk of developing CKD and focus preventative treatment at those at highest risk.

8.8 Clinical implication

8.8.1 Practice

This thesis provides reliable estimates of the incidence of CKD in young adults aged 18 to 40 years with IGR compared to those with normoglycaemia. The unadjusted incidence of CKD was four times higher in IGR patients than individuals with normoglycaemia. After adjustment for potential confounders risk was reduced to approximately three times. These results provide a twofold opportunity to monitor individuals with IGR in primary care by assessing for worsening glycaemic status and monitor kidney function at least annually. Those individuals with IGR and hypertension should be monitored more closely for CKD development. Furthermore, risk of developing CKD stage (3 – 5) was higher than risk of CKD stage (1 – 2) in the IGR cohort, providing further opportunity for individuals with IGR to consider lifestyle modification and have regular follow-up check of their kidney function to detect CKD at an early stage. Among the comorbidities it is worth noting that a proportion of individuals with IGR at baseline already had cardiovascular related complication (1.8%) and hypertension (10.5%). These findings

reinforce the need to consider appropriate preventative treatment (for e.g., anti-hypertensive and cholesterol lowering treatments) along with lifestyle advice following IGR diagnosis.

8.8.2 Policy

The current NICE guidelines on the early identification and management of CKD in adults recommend that patients at risk of developing CKD (e.g. diabetes/hypertension) are offered screening test at least annually to detect further decline in kidney function (9). This study shows that risk of CKD is approximately four times higher in patients with IGR than those with normoglycaemia. Risk of CKD is already elevated even before patient progresses to T2DM. This may suggest that targeted routine screening could potentially be beneficial at IGR diagnosis and initiation of early CKD treatment may slow the decline of kidney function and reduce the risk of kidney failure (9). Furthermore, NICE recommend that patients on prolong treatment with an NSAID should have their kidney function checked at least annually to monitor further decline in GFR (9). In this study 1.9% (200), patients with IGR at baseline were prescribed an NSAID. This provides further opportunity to monitor and screen these patients at an early stage, especially in patients with other risk factors for CKD progression.

8.9 Conclusion

The National Institute for Health and Care Excellence (NICE) recently published guidelines on early identification and management of people diagnosed with CKD. This was driven by Kidney Disease Improving Global Outcomes (KDIGO) current CKD classification and data from recently published prognostic studies. Furthermore, NICE recommend that people with IGR are monitored to prevent future risk of diabetes. This thesis provides evidence of an increased risk of CKD amongst young adults with IGR. Analyses revealed that patients with IGR are at higher risk of developing CKD stage (3 – 5) than CKD stage (1 – 2). Additionally, among the

modifiable risk factors, hypertension was consistently linked to higher incidence of CKD. These results are in agreement with NICE guidelines, that patients at high risk should be tested for CKD. The result of this thesis shows the sizeable adverse effect of IGR on the development of CKD, and if left undetected may cause serious long term health issues and should be addressed at an early stage to reduce the future burden of CKD in patients with IGR. Estimates of the present study may provide a snapshot of young adults aged 18 to 40 years diagnosed with IGR in the UK between the January 2000 and February 2015 and may serve as a baseline comparison with other studies using similar primary care databases in the UK.

9.0 References

1. Stevens PE, Levin A. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. Ann Intern Med 2013;158(11):825-30.

2. Couser WG, Remuzzi G, Mendis S, Tonelli M. The contribution of chronic kidney disease to the global burden of major noncommunicable diseases. Kidney Int 2011;80(12):1258-70.

3. National Institute for Health and Care Excellence (NICE). Quality standard 5: chronic kidney disease 2011. Available from: http://guidance.nice.org.uk/ [Accessed 15th March 2016].

4. Nathan DM, Davidson MB, DeFronzo RA, Heine RJ, Henry RR, Pratley R, et al. Impaired Fasting Glucose and Impaired Glucose Tolerance: Implications for care. Diabetes Care 2007;30(3):753-9.

5. Fox CS, Larson MG, Leip EP, Meigs JB, Wilson PW, Levy D. Glycemic status and development of kidney disease: the Framingham Heart Study. Diabetes Care 2005;28(10):2436-40.

6. Sun F, Tao Q, Zhan S. Metabolic syndrome and the development of chronic kidney disease among 118 924 non-diabetic Taiwanese in a retrospective cohort. Nephrology (Carlton) 2010;15(1):84-92.

7. Watanabe H, Obata H, Watanabe T, Sasaki S, Nagai K, Aizawa Y. Metabolic syndrome and risk of development of chronic kidney disease: the Niigata preventive medicine study. Diabetes Metab Res Rev 2010;26(1):26-32.

8. Chapter 1: Introduction and definition of CKD-MBD and the development of the guideline statements. Kidney Int 2009;76113:S3-8.

9. National Institute for Health and Care Excellence (NICE). Chronic kidney disease in adults: assessment and management 2014. Available from: http://www.nice.org.uk/guidance/cg182 [Accessed 15th March 2016].

10. Kidney Disease Outcome Quality Initiative (K/DOQI) clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis 2002;39(2 Suppl 1):S1-266.

11. Levey AS, Eckardt KU, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int 2005;67(6):2089-100.

12. Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, de Jong PE, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and

cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet 2010;375(9731):2073-81.

13. Astor BC, Matsushita K, Gansevoort RT, van der Velde M, Woodward M, Levey AS, et al. Lower estimated glomerular filtration rate and higher albuminuria are associated with mortality and end-stage renal disease. A collaborative meta-analysis of kidney disease population cohorts. Kidney Int 2011;79(12):1331-40.

14. Gansevoort RT, Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, et al. Lower estimated GFR and higher albuminuria are associated with adverse kidney outcomes in both general and high-risk populations: A collaborative meta-analysis of general and high-risk population cohorts. Kidney Int 2011;80(1):93-104.

15. Levey AS, de Jong PE, Coresh J, El Nahas M, Astor BC, Matsushita K, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO Controversies Conference report. Kidney Int 2011;80(1):17-28.

16. Stevens LA, Coresh J, Greene T, Levey AS. Assessing kidney function--measured and estimated glomerular filtration rate. N Engl J Med 2006;354(23):2473-83.

17. Hojs R, Bevc S, Ekart R, Gorenjak M, Puklavec L. Serum cystatin C as an endogenous marker of renal function in patients with mild to moderate impairment of kidney function. Nephrol Dial Transpl 2006;21(7):1855-62.

18. Baumgarten M, Gehr T. Chronic kidney disease: detection and evaluation. Am Fam Physician 2011;84(10):1138-48.

19. Burden R, Tomson C. Identification, management and referral of adults with chronic kidney disease: concise guidelines. Clin Med (London). 2005;5(6):635-42.

20. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron 1976;16(1):31-41.

21. Levey AS, Perrone RD, Madias NE. Serum creatinine and renal function. Ann Rev Med 1988;39:465-90.

22. Robert S, Zarowitz BJ, Peterson EL, Dumler F. Predictability of creatinine clearance estimates in critically ill patients. Crit Care Med 1993;21(10):1487-95.

23. Swan SK. The search continues--an ideal marker of GFR. Clin Chem 1997;43(6 Pt 1):913-4.

24. Rule AD, Bailey KR, Schwartz GL, Khosla S, Lieske JC, Melton LJ, 3rd. For estimating creatinine clearance measuring muscle mass gives better results than those based on demographics. Kidney Int 2009;75(10):1071-8.

25. Newman DJ, Thakkar H, Edwards RG, Wilkie M, White T, Grubb AO, et al. Serum cystatin C measured by automated immunoassay: a more sensitive marker of changes in GFR than serum creatinine. Kidney Int 1995;47(1):312-8.

26. Kyhse-Andersen J, Schmidt C, Nordin G, Andersson B, Nilsson-Ehle P, Lindstrom V, et al. Serum cystatin C, determined by a rapid, automated particle-enhanced turbidimetric method, is a better marker than serum creatinine for glomerular filtration rate. Clin Chem 1994;40(10):1921-6.

27. Grubb A, Simonsen O, Sturfelt G, Truedsson L, Thysell H. Serum Concentration of Cystatin C, Factor D and β2-Microglobulin as a Measure of Glomerular Filtration Rate. Acta Medica Scandinavica 1985;218(5):499-503.

28. Randers E, Krue S, Erlandsen EJ, Danielsen H, Hansen LG. Reference Interval for Serum Cystatin C in Children. Clin Chem 1999;45(10):1856-8.

29. Fliser D, Ritz E. Serum cystatin C concentration as a marker of renal dysfunction in the elderly. Am J Kidney Dis 2001;37(1):79-83.

30. Lamb E, Brettell E, Cockwell P, Dalton N, Deeks J, Harris K, et al. The eGFR-C study: accuracy of glomerular filtration rate (GFR) estimation using creatinine and cystatin C and albuminuria for monitoring disease progression in patients with stage 3 chronic kidney disease - prospective longitudinal study in a multiethnic population. BMC Nephrol 2014;15(1):13.

31. Taal MW, Brenner BM. Renal risk scores: progress and prospects. Kidney Int 2008;73(11):1216-9.

32. Verhave JC, Hillege HL, Burgerhof JG, Gansevoort RT, de Zeeuw D, de Jong PE. The association between atherosclerotic risk factors and renal function in the general population. Kidney Int 2005;67(5):1967-73.

33. Ishani A, Grandits GA, Grimm RH, Svendsen KH, Collins AJ, Prineas RJ, et al. Association of single measurements of dipstick proteinuria, estimated glomerular filtration rate, and hematocrit with 25-year incidence of end-stage renal disease in the multiple risk factor intervention trial. J Am Soc Nephrol 2006;17(5):1444-52.

34. Viazzi F, Leoncini G, Conti N, Tomolillo C, Giachero G, Vercelli M, et al. Microalbuminuria Is a Predictor of Chronic Renal Insufficiency in Patients without Diabetes and with Hypertension: The MAGIC Study. Clin J Am Soc Nephrol 2010;5(6):1099-106.

35. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. Ann Intern Med 1999;130(6):461-70.

36. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med 2006;145(4):247-54.

37. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150(9):604-12.

38. Poggio ED, Wang X, Greene T, Van Lente F, Hall PM. Performance of the modification of diet in renal disease and Cockcroft-Gault equations in the estimation of GFR in health and in chronic kidney disease. J Am Soc Nephrol 2005;16(2):459-66.

39. Levey AS, Stevens LA. Estimating GFR Using the CKD Epidemiology Collaboration (CKD-EPI) Creatinine Equation: More Accurate GFR Estimates, Lower CKD Prevalence Estimates, and Better Risk Predictions. Am J Kidney Dis 2010;55(4):622-7.

40. Ma YC, Zuo L, Chen JH, Luo Q, Yu XQ, Li Y, et al. Improved GFR estimation by combined creatinine and cystatin C measurements. Kidney Int 2007;72(12):1535-42.

41. Vaidya VS, Ozer JS, Dieterle F, Collings FB, Ramirez V, Troth S, et al. Kidney injury molecule-1 outperforms traditional biomarkers of kidney injury in preclinical biomarker qualification studies. Nature biotechnology 2010;28(5):478-85.

42. Graig R, Mindell J. Health Survey for England 2010. Health Survey For England, 2010. Available from: http://content.digital.nhs.uk/pubs/hse10report [Accessed 26th August 2016].

43. Stevens PE, O'Donoghue DJ, de Lusignan S, Van Vlymen J, Klebe B, Middleton R, et al. Chronic kidney disease management in the United Kingdom: NEOERICA project results. Kidney Int 2007;72(1):92-9.

44. The Information Centre for Health and Social Care. Quality and Outcomes Framework (QOF) Achievement and Prevalence Bulletin 2008-2009: Health and Social Care Information Centre; 2009. Available from: http://www.ic.nhs.uk [Accessed 20th April 2016].

45. The Information Services Division for Scotland 2013. National Prevalence rates by reister and year (April - March, Scotland (2013/14) 2013. Available from: http://www.isdscotland.org/Health-Topics/General-Practice/Quality-And-Outcomes-Framework/ [Accessed 15th April 2016].

46. General Medical Services Contract. Quality and Outcomes Framework (QOF) for 1st April 2013 - 31st March 2014 2013/14. Available from: http://gov.wales/statistics-and-research/general-medical-services-contract/?skip=1&lang=en [Accessed 20th March 2016].

47. Health and Social Services and Public Safety. Quality and Outcomes Framework (QOF) Achievement Data 2013/14 (Northern Ireland) 2013/14. Available from: http://www.dhsspsni.gov.uk/index/statistics/qof/qof-achievement/qof-practice-13-14.htm [Accessed 20th March 2016]

48. World Kidney Day: Chronic kidney disease. Available from: http://www.worldkidneyday.org/faqs/chronic-kidney-disease/ [Accessed 15th April 2016].

49. Coresh J, Selvin E, Stevens LA, Manzi J, Kusek JW, Eggers P, et al. Prevalence of chronic kidney disease in the United States. JAMA 2007;298(17):2038-47.

50. Zhang L, Zhang P, Wang F, Zuo L, Zhou Y, Shi Y, et al. Prevalence and factors associated with CKD: a population study from Beijing. Am J Kidney Dis 2008;51(3):373-84.

51. Chadban SJ, Briganti EM, Kerr PG, Dunstan DW, Welborn TA, Zimmet PZ, et al. Prevalence of kidney damage in Australian adults: The AusDiab kidney study. J Am Soc Nephrol 2003;14(7 Suppl 2):S131-8.

52. Hallan SI, Dahl K, Oien CM, Grootendorst DC, Aasberg A, Holmen J, et al. Screening strategies for chronic kidney disease in the general population: follow-up of cross sectional health survey. BMJ 2006;333(7577):1047.

53. Collins AJ, Foley RN, Herzog C, Chavers B, Gilbertson D, Ishani A, et al. US Renal Data System 2010 Annual Data Report. Am J Kidney Dis 2011;57(1 Suppl 1):A8, e1-526.

54. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, Jafar TH, Heerspink HJ, Mann JF, et al. Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. Lancet (London) 2013;382(9889):339-52.

55. Baigent C, Burbury K, Wheeler D. Premature cardiovascular disease in chronic renal failure. Lancet (London) 2000;356(9224):147-52.

56. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl Med 2004;351(13):1296-305.

57. Parmar MS. Chronic renal disease. BMJ 2002;325(7355):85-90.

58. Hippisley-Cox J, Coupland C. Predicting the risk of chronic Kidney Disease in men and women in England and Wales: prospective derivation and external validation of the QKidney Scores. BMC Fam Pract 2010;11:49.

59. Dreyer G, Hull S, Aitken Z, Chesser A, Yaqoob MM. The effect of ethnicity on the prevalence of diabetes and associated chronic kidney disease. QJM 2009;102(4):261-9.

60. New JP, Middleton RJ, Klebe B, Farmer CK, de Lusignan S, Stevens PE, et al. Assessing the prevalence, monitoring and management of chronic kidney disease in patients with diabetes compared with those without diabetes in general practice. Diabet Med 2007;24(4):364-9.

61. Classification and Diagnosis of Diabetes Mellitus and Other Categories of Glucose Intolerance. Diabetes 1979;28(12):1039-57.

62. WHO. Report of the Expert Committee on Diabetes: WHO technical report series. Geneva: WHO; 1980.

63. Unwin N, Shaw J, Zimmet P, Alberti KG. Impaired glucose tolerance and impaired fasting glycaemia: the current status on definition and intervention. Diabet Med 2002;19(9):708-23.

64. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 1997;20(7):1183-97.

65. National Institute for Health and Social Care (NICE). Preventing Type 2 Diabetes: Risk Identification and Interventions for Individuals at High Risk (PH 38) 2012. Available from: http://www.nice.org.uk/guidance/ph38 [Accessed 15th March 2016].

66. Diabetes UK Position statement: Early identification of people with, and at high risk of Type 2 Diabetes and intervention for those at high risk2014. Available from: https://www.diabetes.org.uk/Documents/About%20Us/What%20we%20say/diabetes-uk-postition-statement-early-identification-type-2-0914.pdf [Accessed 21st November 2016].

67. World Health Organisation (WHO). Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia [internet] 2006. [cited 2016 Aug 26]. Report of WHO/IDF consultation. Available from: http://www.who.int/diabetes/publications/Definition%20and%20diagnosis%20of%20diabetes_new.pdf

68. American Diabetes Association (ADA) Standards of Medical Care in Diabetes—2010. Diabetes Care 2010;33(Supplement 1):S11-S61.

69. John WG. Use of HbA1c in the diagnosis of diabetes mellitus in the UK. The implementation of World Health Organization guidance 2011. Diabet Med 2012;29(11):1350-7.

70. Davies MJ, Muehlbayer S, Garrick P, McNally PG. Potential impact of a change in the diagnostic criteria for diabetes mellitus on the prevalence of abnormal glucose tolerance in a local community at risk of diabetes: impact of new diagnostic criteria for diabetes mellitus. Diabet Medicine 1999;16(4):343-6.

71. College of American Pathologist (CAP). Participants Summary. Northfield: Illnois, 2005.

72. Yudkin JS, Forrest RD, Jackson CA, Burnett SD, Gould MM. The Prevalence of Diabetes and Impaired Glucose Tolerance in a British Population. Diabetes Care 1993;16(11):1530-.

73. Forrest RD, Jackson CA, Yudkin JS. Glucose intolerance and hypertension in north London: the Islington Diabetes Survey. Diabet Medicine 1986;3(4):338-42.

74. Phillips DI, Goulden P, Syddall HE, Aihie Sayer A, Dennison EM, Martin H, et al. Fetal and infant growth and glucose tolerance in the Hertfordshire Cohort Study: a study of men and women born between 1931 and 1939. Diabetes 2005;54 Suppl 2:S145-50.

75. Wilmot EG, Edwardson CL, Biddle SJ, Gorely T, Henson J, Khunti K, et al. Prevalence of diabetes and impaired glucose metabolism in younger 'at risk' UK adults: insights from the STAND programme of research. Diabet Medicine. 2013;30(6):671-5.

76. Mainous AG, Tanner RJ, Baker R, Zayas CE, Harle CA. Prevalence of prediabetes in England from 2003 to 2011: population-based, cross-sectional study. BMJ 2014;4(6).

77. National Cardiovascular Intelligence Network (NCVIN). NHS Diabetes Prevention Programme (NHS DPP) - Non-diabetic hyperglycaemia. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/456149/Non_diabetic_hyperglycaemia.pdf [Accessed 26th August 2016].

78. Gillett M, Royle P, Snaith A, Scotland G, Poobalan A, Imamura M, et al. Non-pharmacological interventions to reduce the risk of diabetes in people with impaired glucose regulation: a systematic review and economic evaluation. Health Technol Assess 2012;16(33):1-236, iii-iv.

79. Unwin N WD, Gan D, Jacqmain O, Ghyoot G. International Diabetes Federation Diabetes Atlas. Belgium: International Diabetes Federation, 2009.

80. Simmons D, Williams DR, Powell MJ. The Coventry Diabetes Study: prevalence of diabetes and impaired glucose tolerance in Europids and Asians. Q J Med 1991;81(296):1021-30.

81. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Medicine 1998;15(7):539-53.

82. Khunti K, Morris DH, Weston CL, Gray LJ, Webb DR, Davies MJ. Joint prevalence of diabetes, impaired glucose regulation, cardiovascular disease risk and chronic kidney disease in South Asians and White Europeans. PloS One 2013;8(1):e55580.

83. Knowler WC, Narayan KM, Hanson RL, Nelson RG, Bennett PH, Tuomilehto J, et al. Preventing non-insulin-dependent diabetes. Diabetes 1995;44(5):483-8.

84. Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. BMJ 2016;355.

85. The European Diabetes Epidemiology Group (DECODE) Glucose tolerance and cardiovascular mortality: comparison of fasting and 2-hour diagnostic criteria. Arch Intern Med 2001;161(3):397-405.

86. Ford ES, Zhao G, Li C. Pre-diabetes and the risk for cardiovascular disease: a systematic review of the evidence. Journal Am Coll Cardiol 2010;55(13):1310-7.

87. Coutinho M, Gerstein HC, Wang Y, Yusuf S. The relationship between glucose and incident cardiovascular events. A metaregression analysis of published data from 20 studies of 95,783 individuals followed for 12.4 years. Diabetes Care 1999;22(2):233-40.

88. Singleton JR, Smith AG, Russell JW, Feldman EL. Microvascular complications of impaired glucose tolerance. Diabetes 2003;52(12):2867-73.

89. Nagi DK, Pettitt DJ, Bennett PH, Klein R, Knowler WC. Diabetic retinopathy assessed by fundus photography in Pima Indians with impaired glucose tolerance and NIDDM. Diabet Medicine 1997;14(6):449-56.

90. Diabetes Prevention Program Research Group. The prevalence of retinopathy in impaired glucose tolerance and recent-onset diabetes in the Diabetes Prevention Program. Diabet Medicine 2007;24(2):137-44.

91. Lee CC, Perkins BA, Kayaniyil S, Harris SB, Retnakaran R, Gerstein HC, et al. Peripheral Neuropathy and Nerve Dysfunction in Individuals at High Risk for Type 2 Diabetes: The PROMISE Cohort. Diabetes Care 2015;38(5):793-800.

92. Metcalf PA, Baker JR, Scragg RKR, Dryson E, Scott AJ, Wild CJ. Microalbuminuria in a Middle-Aged Workforce: Effect of hyperglycemia and ethnicity. Diabetes Care 1993;16(11):1485-93.

93. American Diabetes Association (ADA). Standards of Medical Care in Diabetes—2008. Diabetes Care 2008;31(Supplement 1):S12-S54.

94. Keen H, Jarrett RJ, McCartney P. The ten-year follow-up of the Bedford survey (1962-1972): glucose tolerance and diabetes. Diabetologia 1982;22(2):73-8.

95. Motala AA, Omar MA, Gouws E. High risk of progression to NIDDM in South-African Indians with impaired glucose tolerance. Diabetes 1993;42(4):556-63.

96. de Vegt F, Dekker JM, Jager A, Hienkens E, Kostense PJ, Stehouwer CD, et al. Relation of impaired fasting and postload glucose with incident type 2 diabetes in a Dutch population: The Hoorn Study. JAMA 2001;285(16):2109-13.

97. Gerstein HC, Santaguida P, Raina P, Morrison KM, Balion C, Hunt D, et al. Annual incidence and relative risk of diabetes in people with various categories of dysglycemia: a systematic overview and meta-analysis of prospective studies. Diabetes Res Clin Pract 2007;78(3):305-12.

98. Diaz-Redondo A, Giraldez-Garcia C, Carrillo L, Serrano R, Garcia-Soidan FJ, Artola S, et al. Modifiable risk factors associated with prediabetes in men and women: a cross-sectional analysis of the cohort study in primary health care on the evolution of patients with prediabetes (PREDAPS-Study). BMC Fam Pract 2015;16:5.

99. Morales PA, Mitchell BD, Valdez RA, Hazuda HP, Stern MP, Haffner SM. Incidence of NIDDM and impaired glucose tolerance in hypertensive subjects. The San Antonio Heart Study. Diabetes 1993;42(1):154-61.

100. Tapp RJ, Shaw JE, Zimmet PZ, Balkau B, Chadban SJ, Tonkin AM, et al. Albuminuria is evident in the early stages of diabetes onset: results from the Australian Diabetes, Obesity, and Lifestyle Study (AusDiab). Am J Kidney Dis 2004;44(5):792-8.

101. Plantinga LC, Crews DC, Coresh J, Miller ER, 3rd, Saran R, Yee J, et al. Prevalence of chronic kidney disease in US adults with undiagnosed diabetes or prediabetes. Clin J Am Soc Nephrol 2010;5(4):673-82.

102. Nelson RG, Bennett PH, Beck GJ, Tan M, Knowler WC, Mitch WE, et al. Development and progression of renal disease in Pima Indians with non-insulin-dependent diabetes mellitus. Diabetic Renal Disease Study Group. N Engl J Med 1996;335(22):1636-42.

103. Hoehner CM, Greenlund KJ, Rith-Najarian S, Casper ML, McClellan WM. Association of the insulin resistance syndrome and microalbuminuria among nondiabetic native Americans. The Inter-Tribal Heart Project. J Am Soc Nephrol 2002;13(6):1626-34.

104. NHS England Care Quality Commission, Health Education England, Monitor, Public Health England, Trust Development Authority (2014). NHS Five Year Forward View. London: NHS England. Available from: www.england.nhs.uk/ourwork/futurenhs [Accessed 26th August 2016]

105. Hsiao CJ, Hing E, Socey TC, Cai B. Electronic health record systems and intent to apply for meaningful use incentives among office-based physician practices: United States, 2001-2011. NCHS data brief 2011(79):1-8.

106. Kushniruk AW, Bates DW, Bainbridge M, Househ MS, Borycki EM. National efforts to improve health information system safety in Canada, the United States of America and England. Int J Med Inform 2013;82(5):e149-60.

107. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.

108. Barten DJ, Swinkels LC, Dorsman SA, Dekker J, Veenhof C, de Bakker DH. Treatment of hip/knee osteoarthritis in Dutch general practice and physical therapy practice: an observational study. BMC Fam Pract 2015;16:75.

109. Garcia-Gil M, Elorza JM, Banque M, Comas-Cufi M, Blanch J, Ramos R, et al. Linking of primary care records to census data to study the association between socioeconomic status and cancer incidence in Southern Europe: a nation-wide ecological study. PloS One 2014;9(10):e109706.

110. Shephard E, Stapley S, Hamilton W. The use of electronic databases in primary care research. Family Pract 2011;28(4):352-4.

111. House of Common Science and Technology Committee. The Big Data dilemma: Fourth report of session 2015-16. [internet] 2016. [cited 2016 Aug 26] Available from: https://www.publications.parliament.uk/pa/cm201516/cmselect/cmsctech/468/468.pdf [Accessed 20th April 2016].

112. The Farr Institute of Health Informatics Research. Supporting cutting edge health informatics research across the UK. Available from: http://www.farrinstitute.org/about [Accessed 26th April 2016].

113. The Alan Turing Institute 2016. Available from: http://turing.ac.uk [Accessed 20th April 2016]

114. Department for Business Innovation and Skills. The Allocation of Science and Research Funding 2015/16: Investing in World Class Sceince and Research. London: The Stationery Office; 2014 . Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/332767/bis-14-750-science-research-funding-allocations-2015-2016-corrected.pdf [Accessed 15th March 2016].

115. The Health Improvement Network (THIN) 2016. Available from: http://www.thin-uk.com/partnership.htm [Accessed 20th April 2016].

116. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiology. 2015.

117. QRESEARCH. QRESEARCH Practice Information Sheet 2016. Available from: http://www.qresearch.org/SitePages/Information%20for%20practices.aspx [Accessed 18th January 2016].

118. Gill P, Foskett-Tharby R, Hex N. Pay-for-performance and primary care physicians: lessons from the UK Quality and Outcomes Framework for local incentive schemes. J R Soc Med 2015;108(3):80-2.

119. Boyle S, Appleby J, Harrison A. A Rapid View of Access to Care: The Kings Fund 2010. Available from: https://www.kingsfund.org.uk/sites/files/kf/field/field_document/rapid-view-access-care-gpinquiry-research-paper-mar11.pdf [Accessed 22nd January 2016].

120. Carey IM, Cook DG, De Wilde S, Bremner SA, Richards N, Caine S, et al. Implications of the problem orientated medical record (POMR) for research using electronic GP databases: a comparison of the Doctors Independent Network Database (DIN) and the General Practice Research Database (GPRD). BMC Fam Pract 2003;4:14.

121. Shephard E, Stapley S, Hamilton W. The use of electronic databases in primary care research. Fam Pract 2011;28(4):352-4.

122. ResearchOne. ResearchOne 2016. Available from: http://www.researchone.org/ [Accessed 22nd January 2016].

123. Cegedim Medical Research UK. THIN Data Guide for Researchers v2.9. Cegedim Medical Research UK; 2014.

124. Thomas KH, Davies N, Metcalfe C, Windmeijer F, Martin RM, Gunnell D. Validation of suicide and self-harm records in the Clinical Practice Research Datalink. Bri J Clin Pharmacol 2013;76(1):145-57.

125. University College London. THIN data: Strengths and Limitations 2013. Available from: www.ucl.ac.uk/pcph/research-groups-themes/thin-pub/database/pros-cons [Accessed 18th January 2016].

126. Health and Social Care Information Centre Statistics. Available from: http://www.hscic.gov.uk/hes[Accessed 8th March 2016]

127. IMS Health Intelligence Applied. About Us 2015. Available from: http://www.csdmruk.imshealth.com/about-us/about-us.shtml. [Accessed 18th January 2016].

128. Haynes K, Bilker WB, TenHave TR, Strom BL, Lewis JD. Temporal and Within Practice Variability in The Health Improvement Network. Pharmacoepidemiol Drug Saf 2011;20(9):948-55.

129. IMS Health Intelligence Applied. Data Collection 2015. Available from: http://csdmruk.cegedim.com/our-data/data-content.shtml [Accessed 18th January 2016].

130. IMS Health Intelligence Applied. Data Content 2015. Available from: http://csdmruk.cegedim.com/our-data/data-collection.shtml [Accessed 18th January 2016]

131. UK Terminology Centre. Why migrate to SNOMED CT. NHS: London: 2011 Version 1.1.

132. RESIP UK: A Cegedim Company. Vision to use Gemscript 2012. Available from: http://www.resip.co.uk/news/vision-use-gemscript [Accessed 10th February 2016].

133. British National Formulary [BNF] 2014 [14 Jan 2014]. Available from: Available at: http://www.bnf.org/bnf [Accessed 10th February 2016].

134. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. Pharmacoepidemiol Drug Saf 2009;18(1):76-83.

135. Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. BMJ 1991;302(6779):766-8.

136. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Inform Prim Care 2011;19(4):251-5.

137. Denburg MR, Haynes K, Shults J, Lewis JD, Leonard MB. Validation of The Health Improvement Network (THIN) database for epidemiologic studies of chronic kidney disease. Pharmacoepidemiol Drug Saf 2011;20(11):1138-49.

138. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. Bri J Gen Practice 2010;60(572):e128-36.

139. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. Bri J Clin Pharmacol 2010;69(1):4-14.

140. Jordan K, Porcheret M, Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. Fam Pract 2004;21(4):396-412.

141. Anandarajah S, Tai T, de Lusignan S, Stevens P, O'Donoghue D, Walker M, et al. The validity of searching routinely collected general practice computer data to identify patients with chronic kidney disease (CKD): a manual review of 500 medical records. Nephrol Dial Transpl 2005;20(10):2089-96.

142. Health and Social Care Information Centre (hscic). Available from: http://qof.hscic.gov.uk/ [Accessed 11th March 2016].

143. Walker N, Bankart J, Brunskill N, Baker R. Which factors are associated with higher rates of chronic kidney disease recording in primary care? A cross-sectional survey of GP practices. Bri J Gen Pract 2011;61(584):203-5.

144. Bourke A, Dattani H, Robinson M. Feasibility study and methodology to create a quality-evaluated database of primary care data. Inform Prim Care 2004;12(3):171-7.

145. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. Pharmacoepidemiol Drug Saf. 2007;16(4):393-401.

146. Levey AS, Coresh J, Balk E, Kausz AT, Levin A, Steffes MW, et al. National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Ann Intern Med 2003;139(2):137-47.

147. Sechi LA, Catena C, Zingaro L, Melis A, De Marchi S. Abnormalities of glucose metabolism in patients with early renal failure. Diabetes 2002;51(4):1226-32.

148. Kubo M, Kiyohara Y, Kato I, Iwamoto H, Nakayama K, Hirakata H, et al. Effect of hyperinsulinemia on renal function in a general Japanese population: the Hisayama study. Kidney Int 1999;55(6):2450-6.

149. Chen J, Muntner P, Hamm LL, Fonseca V, Batuman V, Whelton PK, et al. Insulin resistance and risk of chronic kidney disease in nondiabetic US adults. J Am Soc Nephrol 2003;14(2):469-77.

150. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009; 339:b2700.

151. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283(15):2008-12.

152. Jadhakhan F, Marshall T, Gill P. A systematic review investigating the cumulative incidence of chronic kidney disease in young adults with impaired glucose tolerance. Syst Rev 2015;4:69.

153. Buysschaert M, Bergman M. Definition of prediabetes. The Medical clinics of North America. 2011;95(2):289-97, vii.

154. American Diabetes Association (ADA) Diagnosis and classification of diabetes mellitus. Diabetes care. 2010;33 Suppl 1:S62-9.

155. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Int Med 2006;144(6):427-37.

156. Hartling L, Milne A, Hamm MP, Vandermeer B, Ansari M, Tsertsvadze A, et al. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. J Clin Epidemiol 2013;66(9):982-93.

157. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

158. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50(4):1088-101.

159. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629-34.

160. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol 2000;53(11):1119-29.

161. Melsen WG, Bootsma MCJ, Rovers MM, Bonten MJM. The effects of clinical and statistical heterogeneity on the predictive values of results from meta-analyses. Clinical Microbiol Infect 2014;20(2):123-9.

162. Meigs JB, D'Agostino RB, Sr., Nathan DM, Rifai N, Wilson PW. Longitudinal association of glycemia and microalbuminuria: the Framingham Offspring Study. Diabetes Care 2002;25(6):977-83.

163. Nelson RG, Tan M, Beck GJ, Bennett PH, Knowler WC, Mitch WE, et al. Changing glomerular filtration with progression from impaired glucose tolerance to Type II diabetes mellitus. Diabetologia 1999;42(1):90-3.

164. Nelson RG, Kunzelman CL, Pettitt DJ, Saad MF, Bennett PH, Knowler WC. Albuminuria in type 2 (non-insulin-dependent) diabetes mellitus and impaired glucose tolerance in Pima Indians. Diabetologia 1989;32(12):870-6.

165. Yokoyama H, Kanno S, Takahashi S, Yamada D, Itoh H, Saito K, et al. Determinants of Decline in Glomerular Filtration Rate in Nonproteinuric Subjects with or without Diabetes and Hypertension. Clin J Am Soc Nephrol 2009;4(9):1432-40.

166. Tozawa M, Iseki C, Tokashiki K, Chinen S, Kohagura K, Kinjo K, et al. Metabolic Syndrome and Risk of Developing Chronic Kidney Disease in Japanese Adults. Hypertens Res 2007;30(10):937-43.

167. Nelson RG, Pettitt DJ, Baird HR, Charles MA, Liu QZ, Bennett PH, et al. Pre-diabetic blood pressure predicts urinary albumin excretion after the onset of type 2 (non-insulin-dependent) diabetes mellitus in Pima Indians. Diabetologia 1993;36(10):998-1001.

168. Rashidi A, Ghanbarian A, Azizi F. Are patients who have metabolic syndrome without diabetes at risk for developing chronic kidney disease? Evidence based on data from a large cohort screening population. Clin J Am Soc Nephrol 2007;2(5):976-83.

169. Kitiyakara C, Yamwong S, Cheepudomwit S, Domrongkitchaiporn S, Unkurapinun N, Pakpeankitvatana V, et al. The metabolic syndrome and chronic kidney disease in a Southeast Asian cohort. Kidney Int 2007;71(7):693-700.

170. Yang T, Chu CH, Hsu CH, Hsieh PC, Chung TC, Bai CH, et al. Impact of metabolic syndrome on the incidence of chronic kidney disease: a Chinese cohort study. Nephrology (Carlton) 2012;17(6):532-8.

171. Kovács T, Vas T, Kovesdy CP, Késõi I, Sági B, Wittmann I, et al. Metabolic syndrome and other cardiovascular risk factors associated with the progression of IgA nephropathy. Clin Kidney J 2012.

172. Ryu S, Chang Y, Woo HY, Lee KB, Kim SG, Kim DI, et al. Time-dependent association between metabolic syndrome and risk of CKD in Korean men without hypertension or diabetes. Am J Kidney Diseases 2009;53(1):59-69.

173. Tohidi M, Hasheminia M, Mohebi R, Khalili D, Hosseinpanah F, Yazdani B, et al. Incidence of chronic kidney disease and its risk factors, results of over 10 year follow up in an Iranian cohort. PloS One 2012;7(9):e45304.

174. Jee SH, Boulware LE, Guallar E, Suh I, Appel LJ, Miller ER, 3rd. Direct, progressive association of cardiovascular risk factors with incident proteinuria: results from the Korea Medical Insurance Corporation (KMIC) study. Arch Intern Med 2005;165(19):2299-304.

175. Carson AP, Muntner P, Carnethon MR, Gross MD, Lewis CE. Abstract P171: Hemoglobin A1c and Incident Albuminuria Among Individuals Without Diabetes: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. Circulation. 2015;131(Suppl 1):AP171-AP.

176. Nand N, Sharma, M., Anand, Lovkesh., Shankar, V., Kumar Hitendra. Evaluation of renal functions in patients having metabolic syndrome in Asian Indian cohort. J Ind Acad Clin Med 2015;16(1):33-8.

177. Halbesma N, Brantsma AH, Bakker SJ, Jansen DF, Stolk RP, De Zeeuw D, et al. Gender differences in predictors of the decline of renal function in the general population. Kidney Int 2008;74(4):505-12.

178. Schottker B, Brenner H, Koenig W, Muller H, Rothenbacher D. Prognostic association of HbA1c and fasting plasma glucose with reduced kidney function in subjects with and without diabetes mellitus. Results from a population-based cohort study from Germany. Prev Med 2013;57(5):596-600.

179. Bonnet F, Marre M, Halimi JM, Stengel B, Lange C, Laville M, et al. Waist circumference and the metabolic syndrome predict the development of elevated albuminuria in non-diabetic subjects: the DESIR Study. J Hypertens 2006;24(6):1157-63.

180. Lucove J, Vupputuri S, Heiss G, North K, Russell M. Metabolic syndrome and the development of CKD in American Indians: the Strong Heart Study. Am J Kidney Dis 2008;51(1):21-8.

181. Ninomiya T, Kiyohara Y, Kubo M, Yonemoto K, Tanizaki Y, Doi Y, et al. Metabolic syndrome and CKD in a general Japanese population: the Hisayama Study. Am J Kidney Dis 2006;48(3):383-91.

182. Kim NH, Pavkov ME, Knowler WC, Hanson RL, Weil EJ, Curtis JM, et al. Predictive value of albuminuria in American Indian youth with or without type 2 diabetes. Pediatrics 2010;125(4):e844-51.

183. Iseki K, Ikemiya Y, Kinjo K, Iseki C, Takishita S. Prevalence of high fasting plasma glucose and risk of developing end-stage renal disease in screened subjects in Okinawa, Japan. Clin Expe Nephrol 2004;8(3):250-6.

184. Echouffo-Tcheugui JB, Narayan KM, Weisman D, Golden SH, Jaar BG. Association between prediabetes and risk of chronic kidney disease: a systematic review and meta-analysis. Diabetic medicine : a journal of the British Diabetic Association. 2016.

185. The Health Improvement Network (THIN) Database. The Health Improvement Network (THIN) Database. Available from: http://csdmruk.cegedim.com [Accessed 15th March 2016].

186. World Health Organisation Guidelines: Approved by the Review Committee. Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus. Abbreviated Report of a WHO Consultation. Geneva: World Health Organisation (2011).

187. Collins GS, Altman DG. An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. BMJ 2009;339.

188. Szatkowski L, Lewis S, McNeill A, Huang Y, Coleman T. Can data from primary care medical records be used to monitor national smoking prevalence? J Epidemiol Community Health 2012;66(9):791-5.

189. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley and Sons; 1987.

190. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med 1991;10(4):585-98.

191. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, Petersen I. Issues in multiple imputation of missing data for large general practice clinical databases. Pharmacoepidemiol Drug Saf 2010;19(6):618-26.

192. STATA. STRATE: Tabulate failure, rates and rate ratios 2016. Available from: http://www.stata.com/manuals13/ststrate.pdf [Accessed 18th November 2016].

193. STATA. Poisson - Poisson regression 2016. Available from: http://www.stata.com/manuals13/rpoisson.pdf [Accessed 18th November 2016].

194. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 1999;28(5):964-74.

195. Cox DR. Regression Models and Life-Tables. J R Stat Soc 1972;34(2):187-220.

196. STATA. stcox - Cox proportional hazards model. Available from: Available from: http://www.stata.com/manuals13/ststcox.pdf [Accessed 31st May 2016].

197. STATA. MFP: Multivariable fractional polynomial models. Available from: http://www.stata.com/manuals13/rmfp.pdf [Accessed 31st May 2016].

198. Rich JT, Neely JG, Paniello RC, Voelker CCJ, Nussenbaum B, Wang EW. A PRACTICAL GUIDE TO UNDERSTANDING KAPLAN-MEIER CURVES. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2010;143(3):331-6.

199. STATA. STS list: List the survivor or cumulative hazards function. Available from: http://www.stata.com/manuals13/ststslist.pdf [Accessed 31st May 2016].

200. STATA. STCOX - PH - assumption tests - Test of proportional - hazards assumption. Available from: http://www.stata.com/manuals13/ststcoxph-assumptiontests.pdf [Accessed 31st May 2016].

201. Collette D. Modelling Survival Data in Medical Research. London; New-York; Washington: Chapman and Hall; 2003.

202. Royston P. Explained variation for survival models. Stata Journal. 2006;6(1):83-96.

203. Nagelkerne NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78(3):691-2.

204. Royston P AG. "Sg81: Multivariable Fractional Polynomials. Stata Technical Bulletin. 1998(43):24-32.

205. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Stat Med 2004;23(5):723-48.

206. ten Ham-Baloyi W, Jordan P. Systematic review as a research method in post-graduate nursing education. Health SA Gesondheid 2016;21:120-8.

207. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339.

208. Long AN, Dagogo-Jack S. Comorbidities of diabetes and hypertension: mechanisms and approach to target organ protection. J Clin Hypertens 2011;13(4):244-51.

209. Retnakaran R, Cull CA, Thorne KI, Adler AI, Holman RR. Risk factors for renal dysfunction in type 2 diabetes: U.K. Prospective Diabetes Study 74. Diabetes 2006;55(6):1832-9.

210. Rao SS, Disraeli P, McGregor T. Impaired glucose tolerance and impaired fasting glucose. Am Fam Physician 2004;69(8):1961-8.

211. Thomas G, Sehgal AR, Kashyap SR, Srinivas TR, Kirwan JP, Navaneethan SD. Metabolic Syndrome and Kidney Disease: A Systematic Review and Meta-analysis. Clin J Am Soc Nephrol 2011;6(10):2364-73.

212. Kurella M, Lo JC, Chertow GM. Metabolic Syndrome and the Risk for Chronic Kidney Disease among Nondiabetic Adults. J Am Soc Nephrol 2005;16(7):2134-40.

213. Anjana RM, Shanthi Rani CS, Deepa M, Pradeepa R, Sudha V, Divya Nair H, et al. Incidence of Diabetes and Prediabetes and Predictors of Progression Among Asian Indians: 10-Year Follow-up of the Chennai Urban Rural Epidemiology Study (CURES). Diabetes Care 2015;38(8):1441-8.

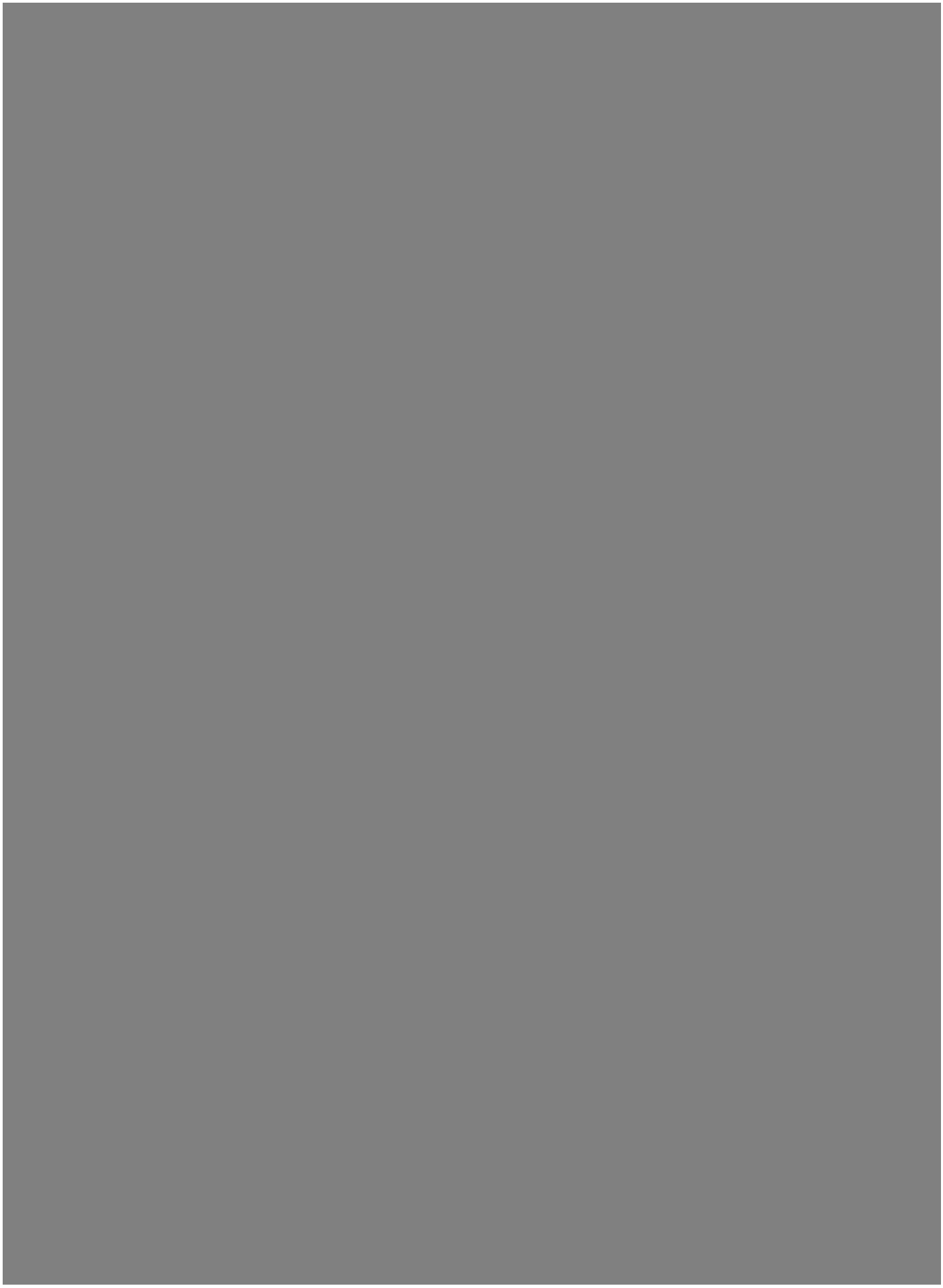
214. Dagogo-Jack S, Edeoga C, Ebenibo S, Nyenwe E, Wan J. Lack of racial disparity in incident prediabetes and glycemic progression among black and white offspring of parents with type 2 diabetes: the pathobiology of prediabetes in a biracial cohort (POP-ABC) study. J Clin Endocrinol Metab 2014;99(6):E1078-87.

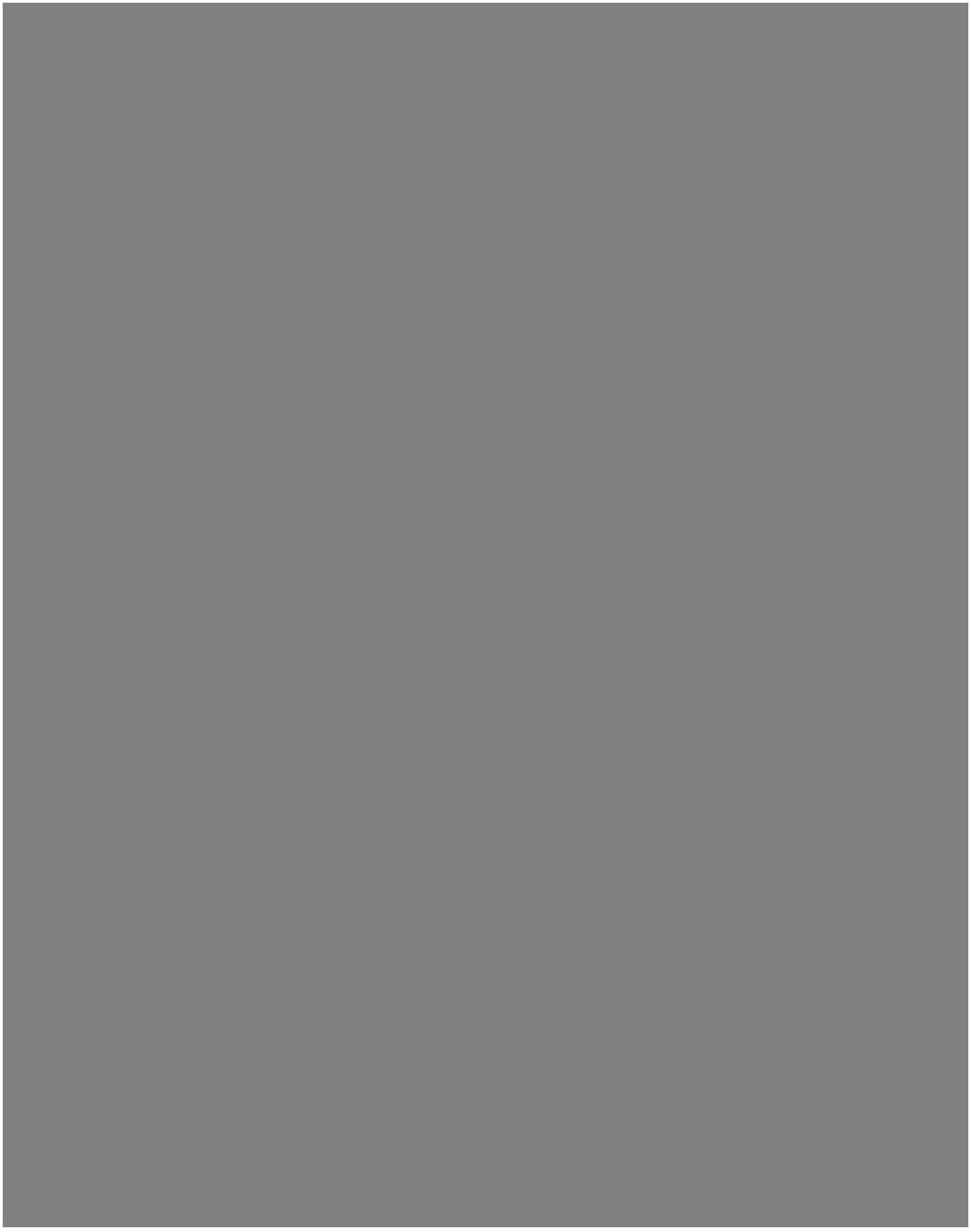
215. Dunstan DW, Zimmet PZ, Welborn TA, de Courten MP, Cameron AJ, Sicree RA, et al. The Rising Prevalence of Diabetes and Impaired Glucose Tolerance. The Australian Diabetes, Obesity and Lifestyle Study 2002;25(5):829-34.

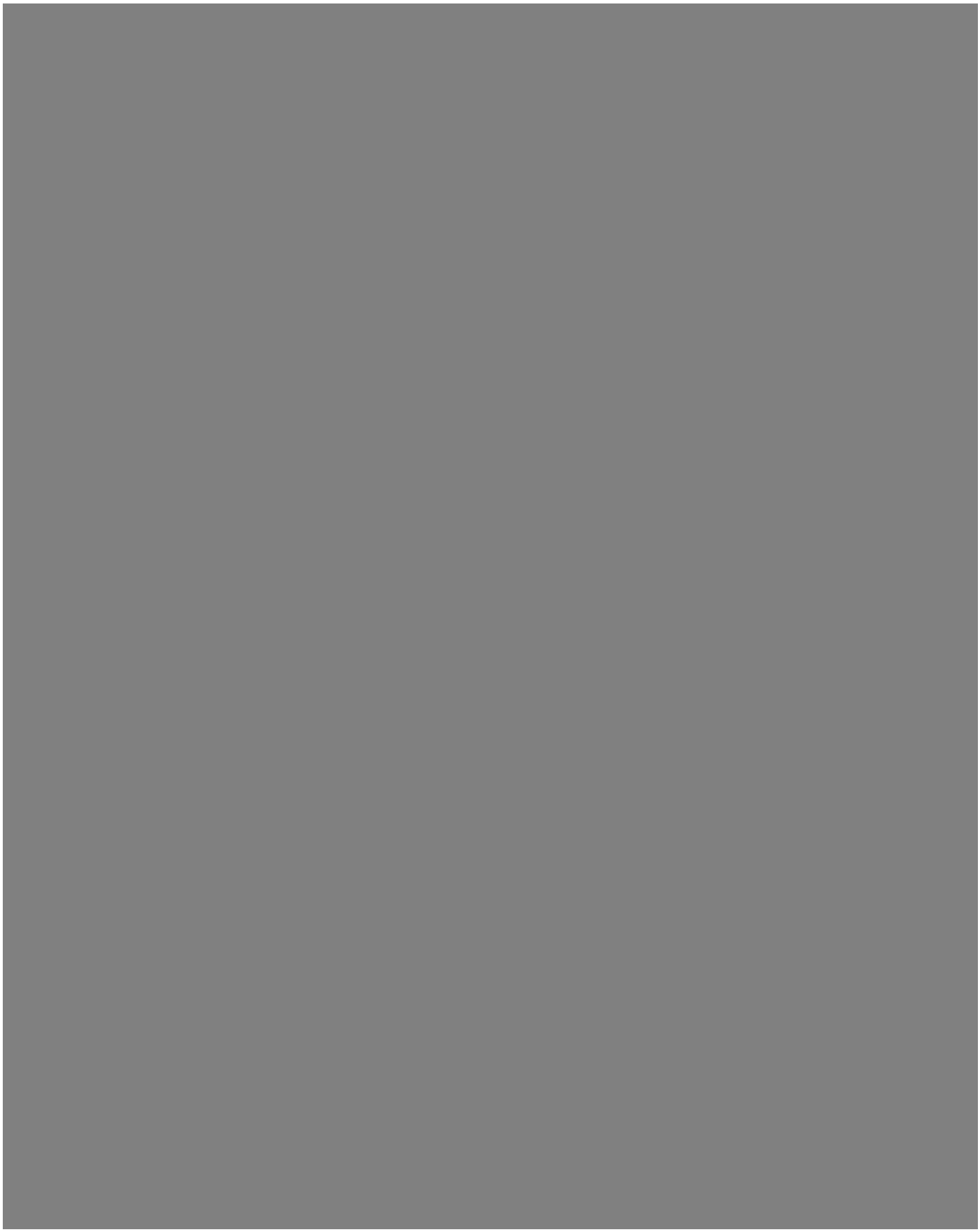
216. Yamagata K, Ishida K, Sairenchi T, Takahashi H, Ohba S, Shiigai T, et al. Risk factors for chronic kidney disease in a community-based population: a 10-year follow-up study. Kidney Int 2007;71(2):159-66.

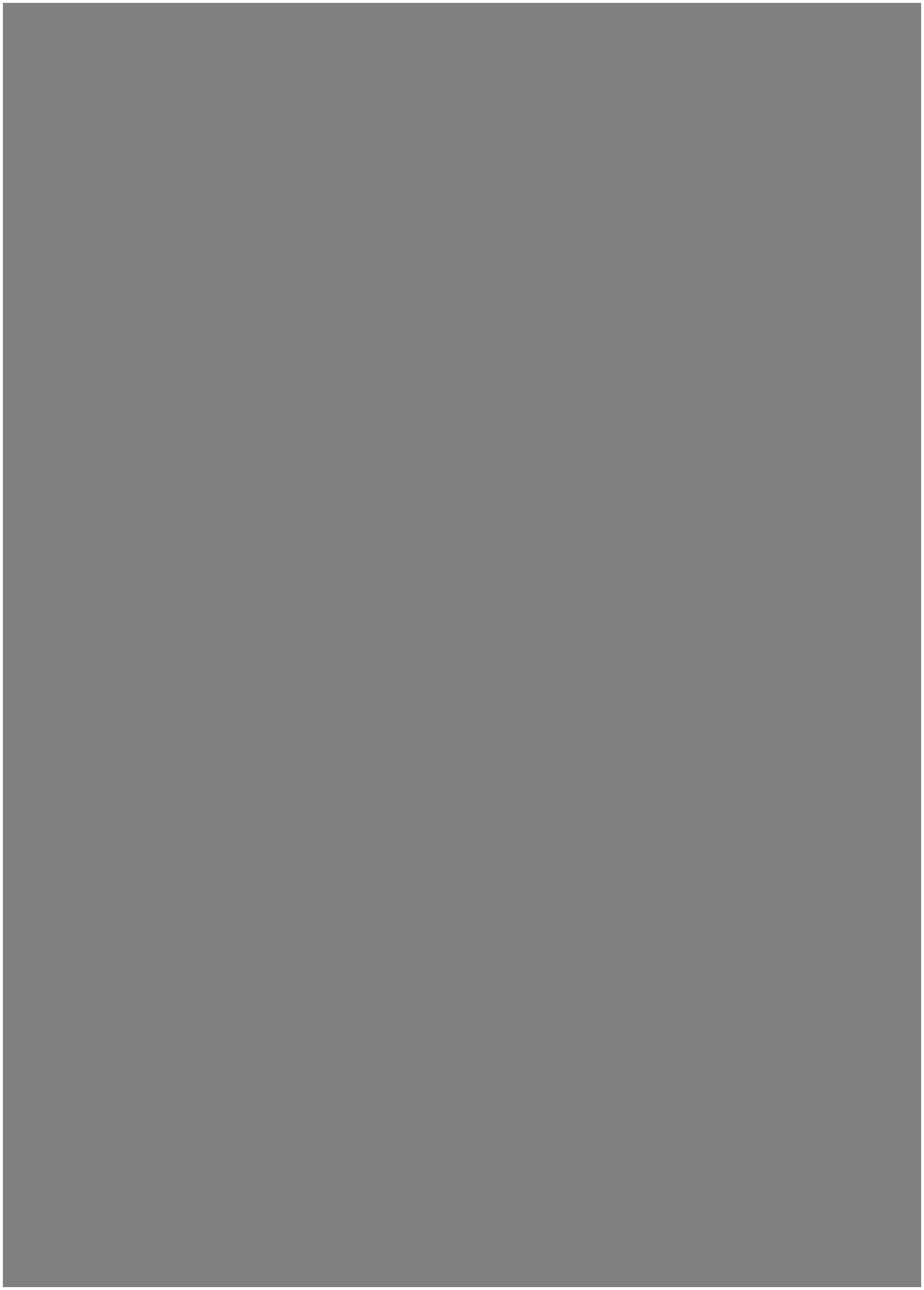
217. Dunkler D, Gao P, Lee SF, Heinze G, Clase CM, Tobe S, et al. Risk Prediction for Early CKD in Type 2 Diabetes. Clin J Am Soc Nephrol 2015.

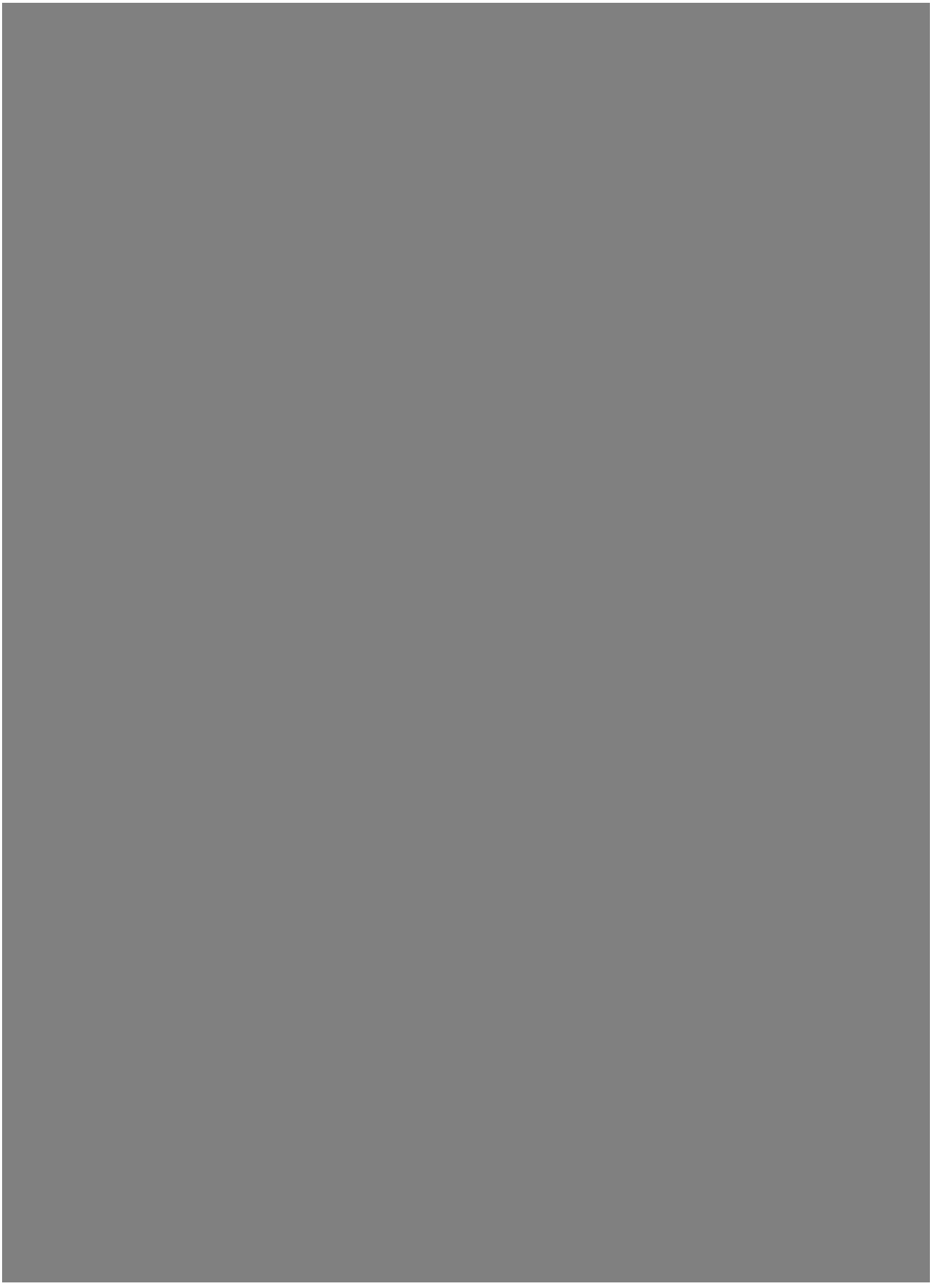
218. Riphagen IJ, Kleefstra N, Drion I, Alkhalaf A, van Diepen M, Cao Q, et al. Comparison of Methods for Renal Risk Prediction in Patients with Type 2 Diabetes (ZODIAC-36). PloS One. 2015;10(3):e0120477.

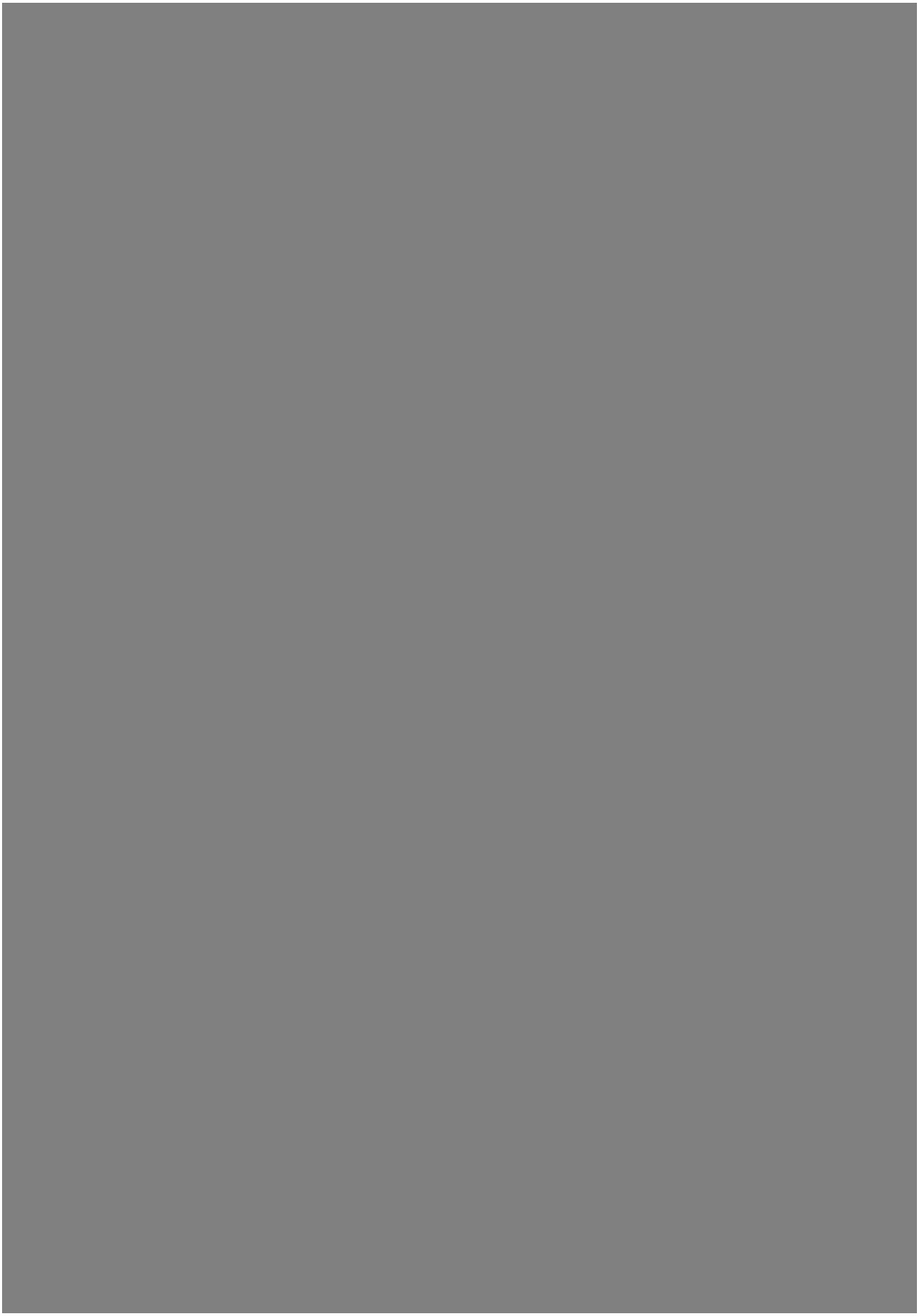


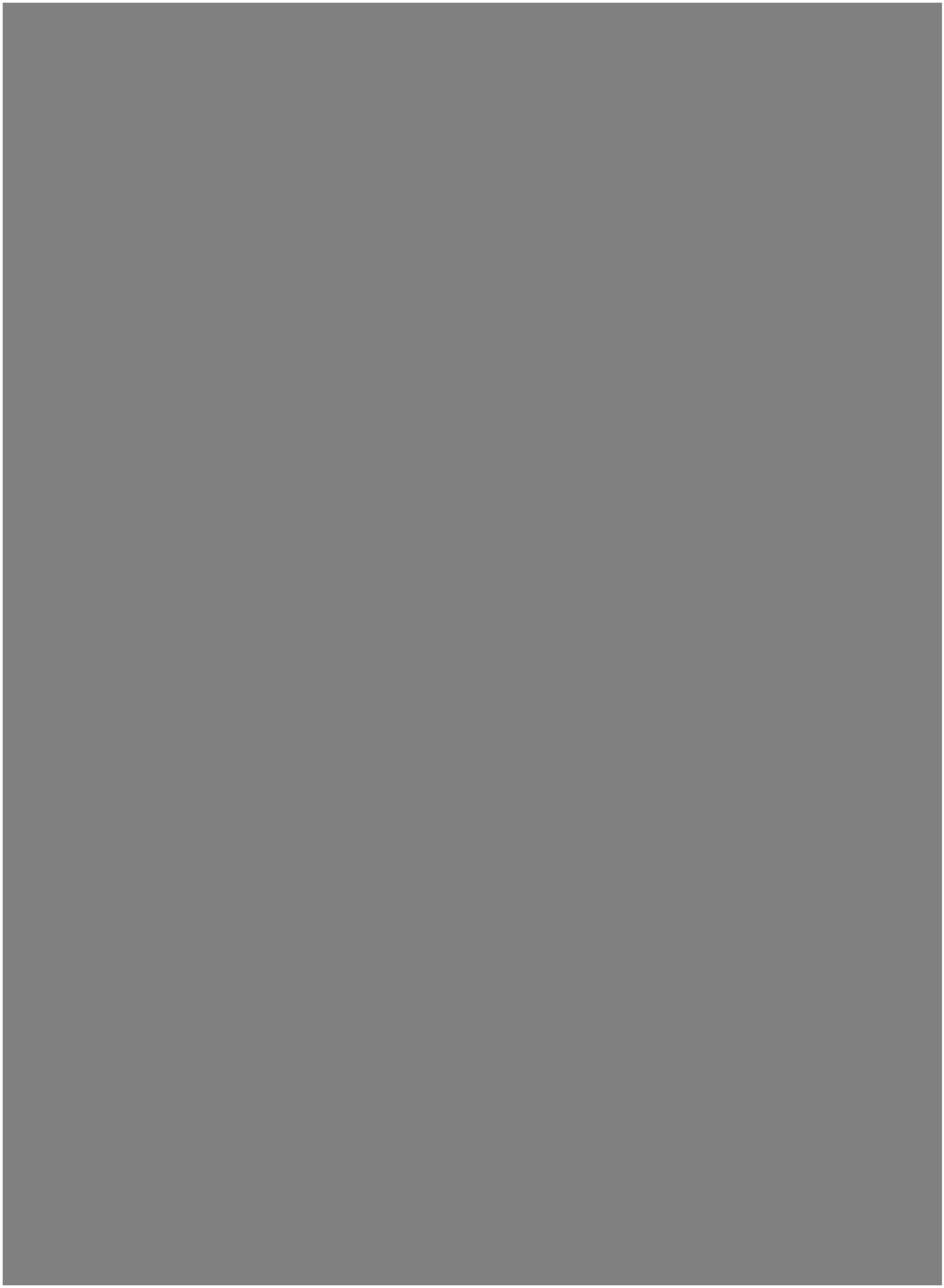


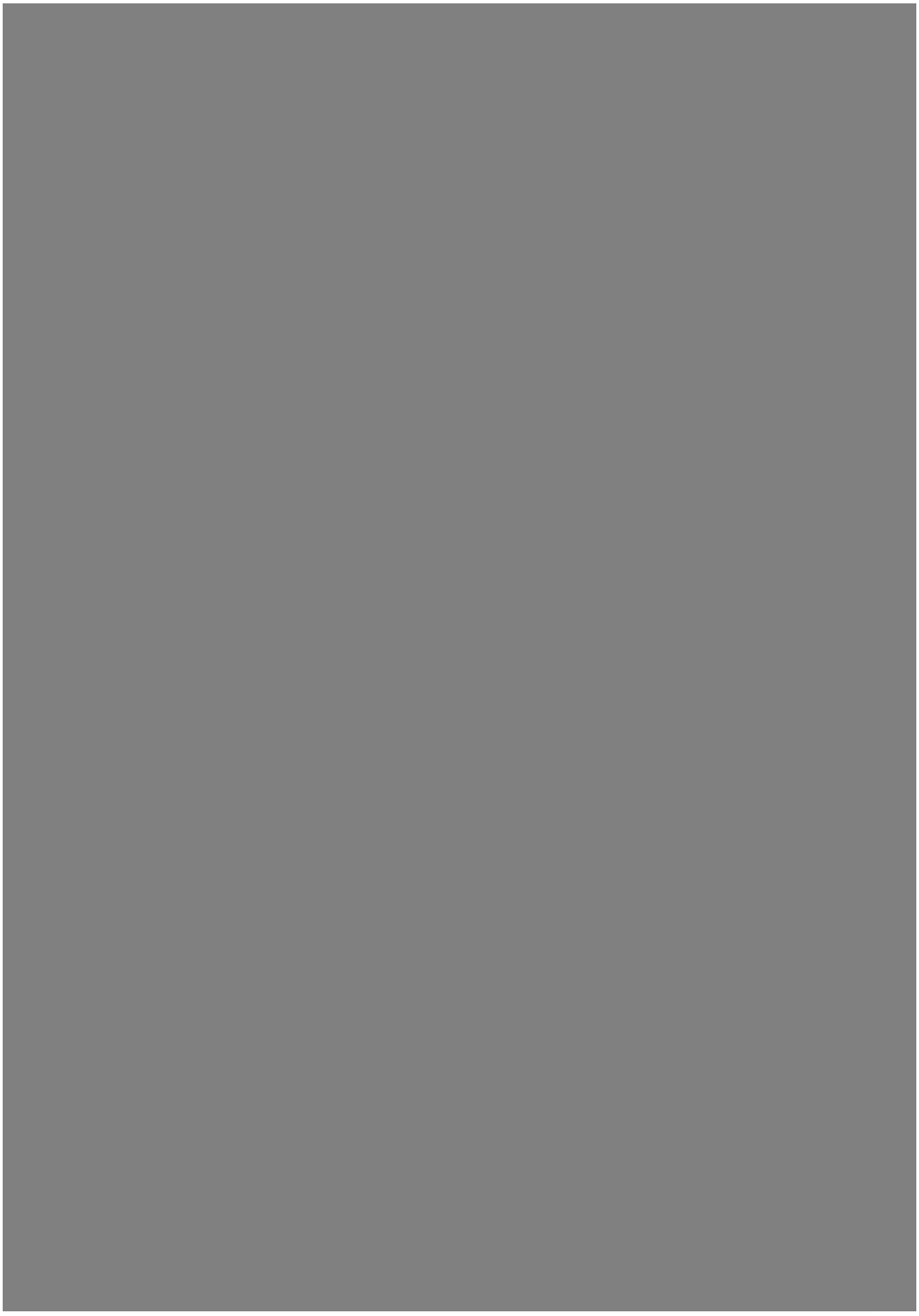


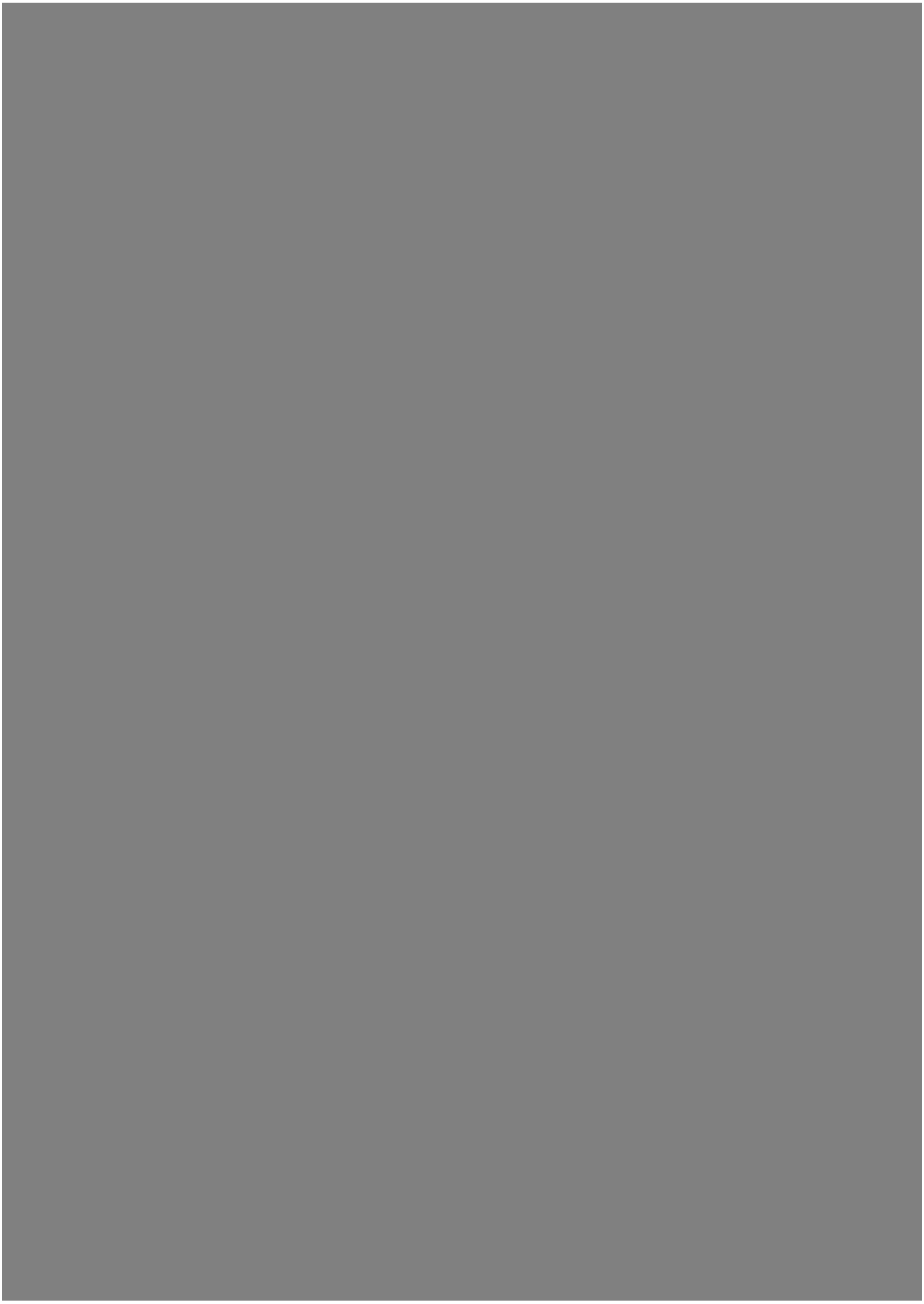


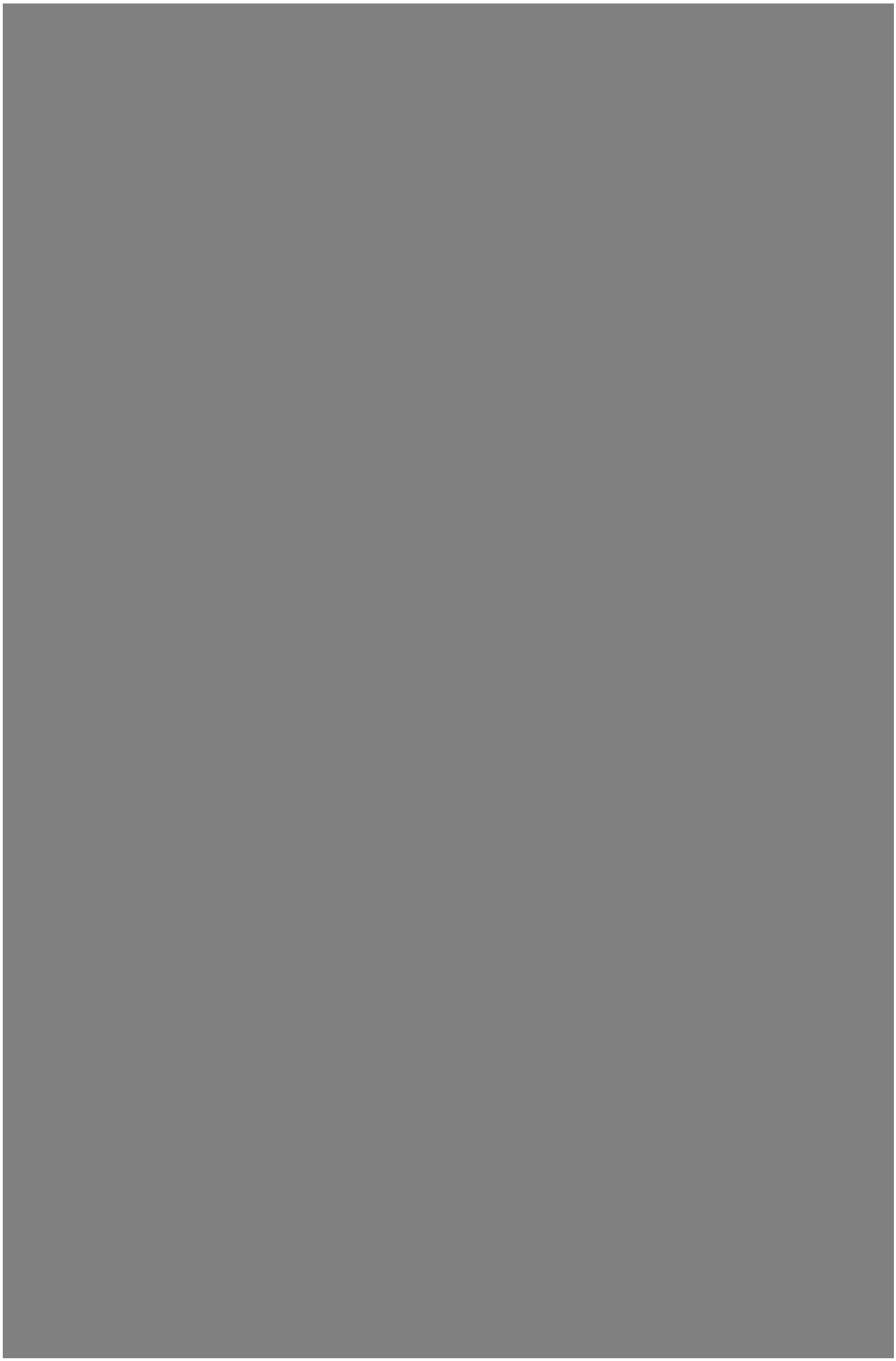


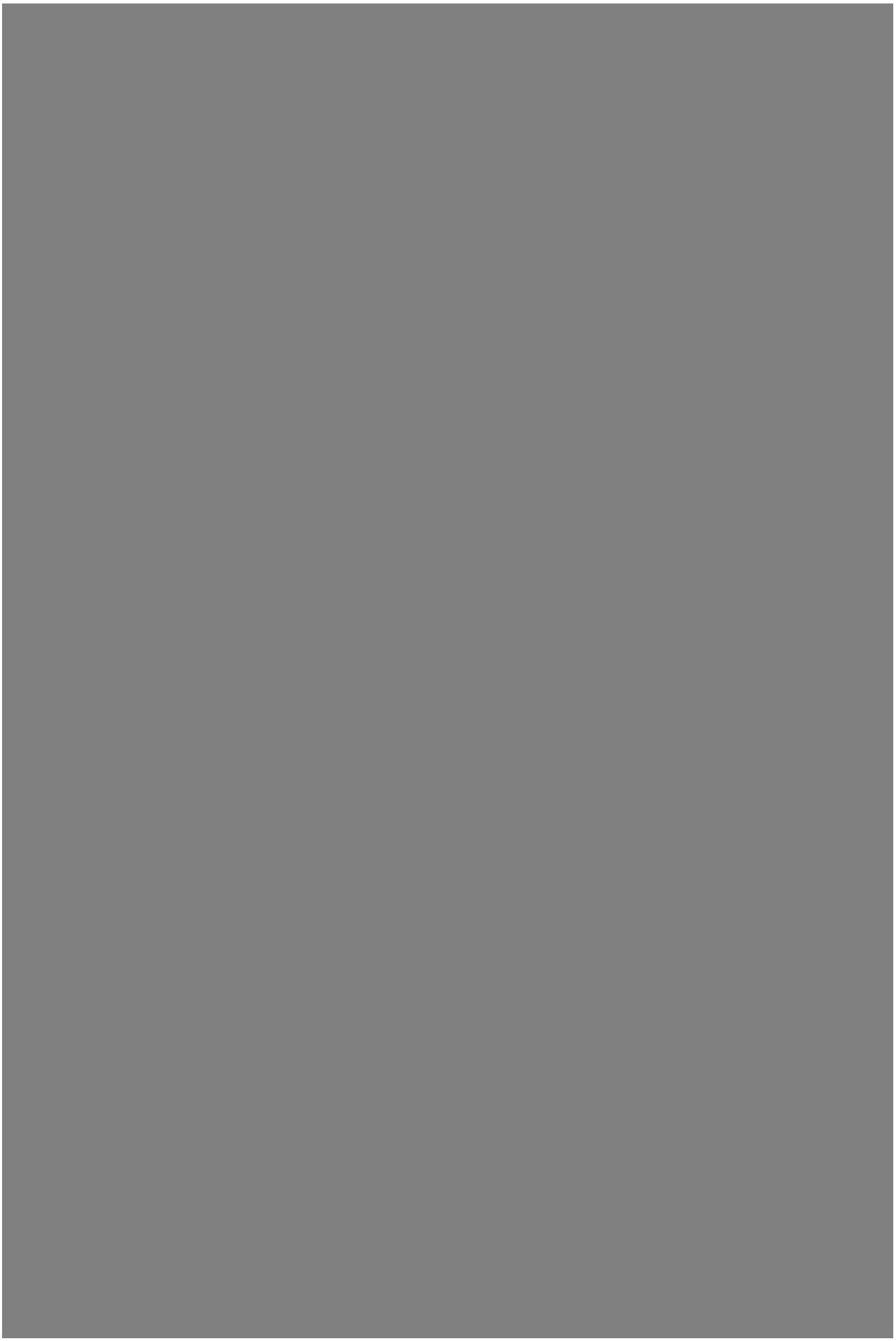


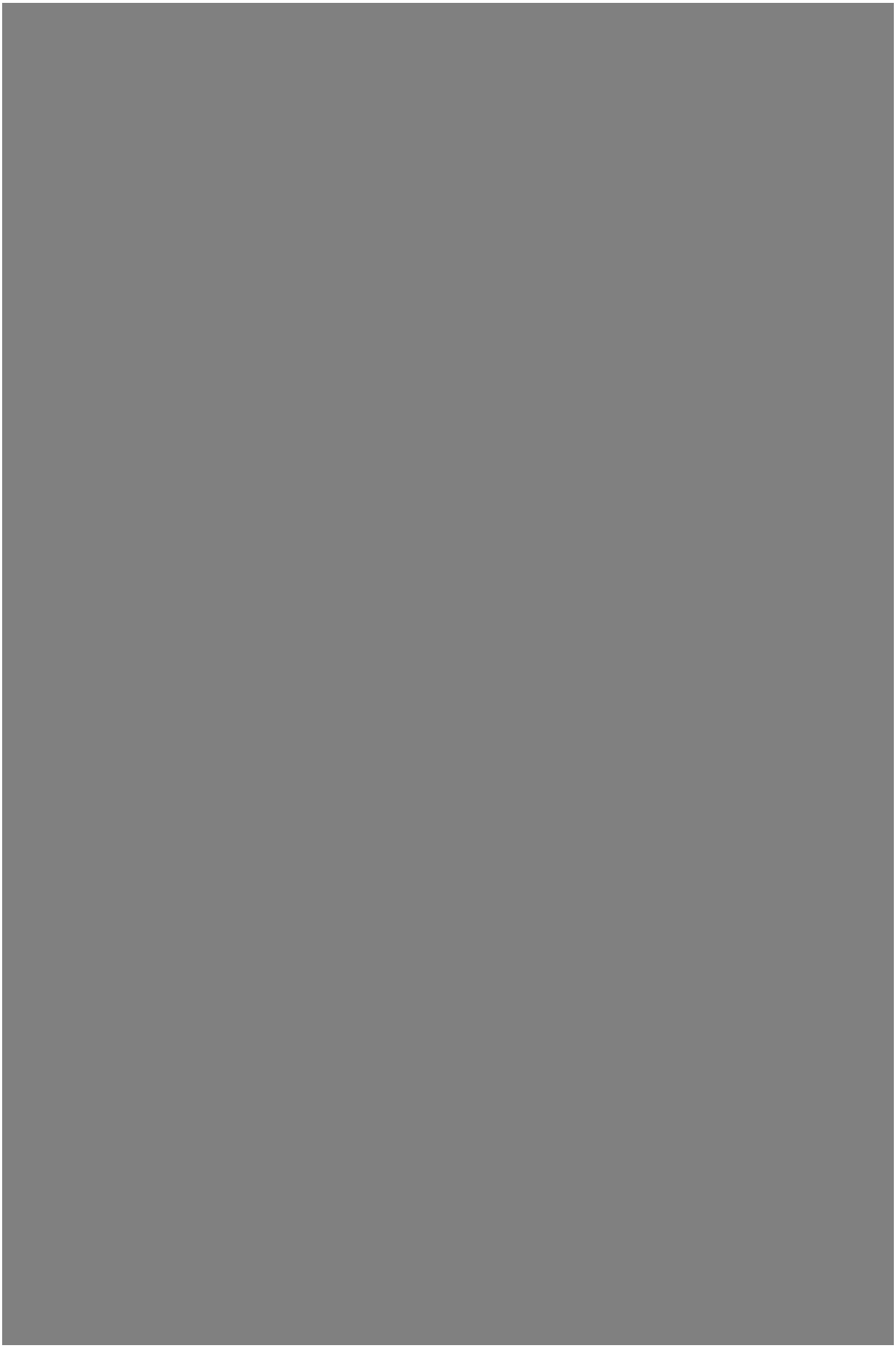


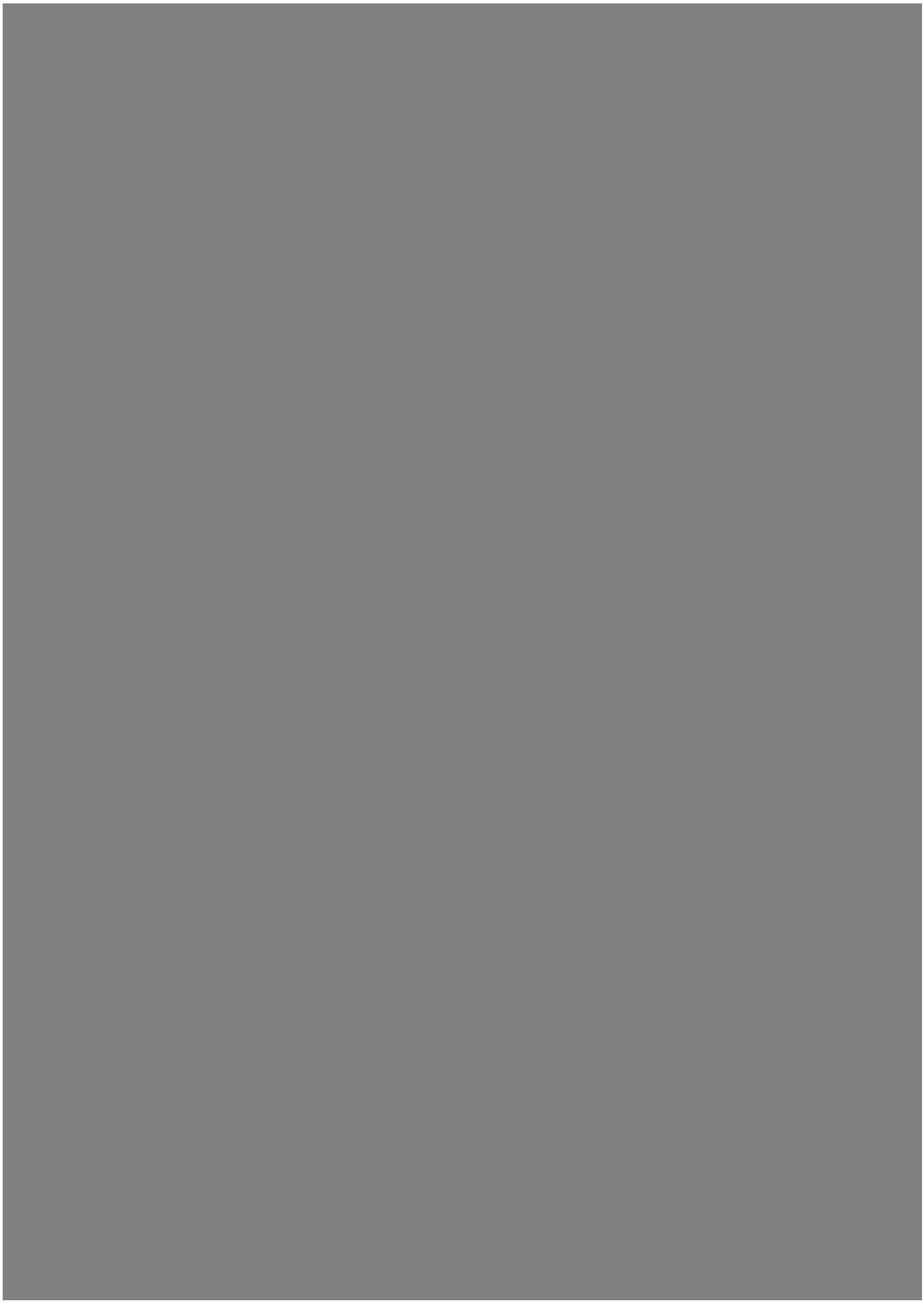


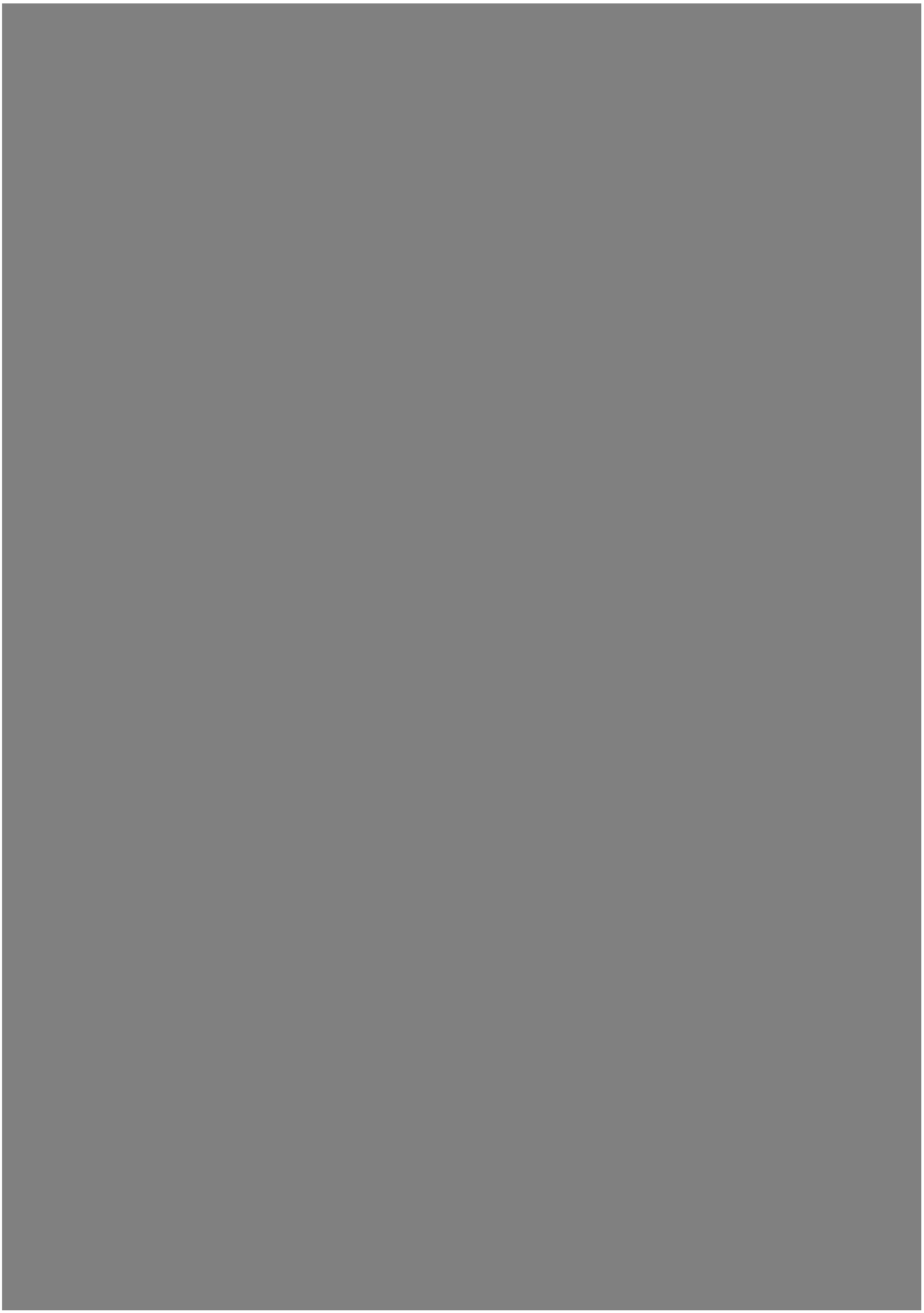


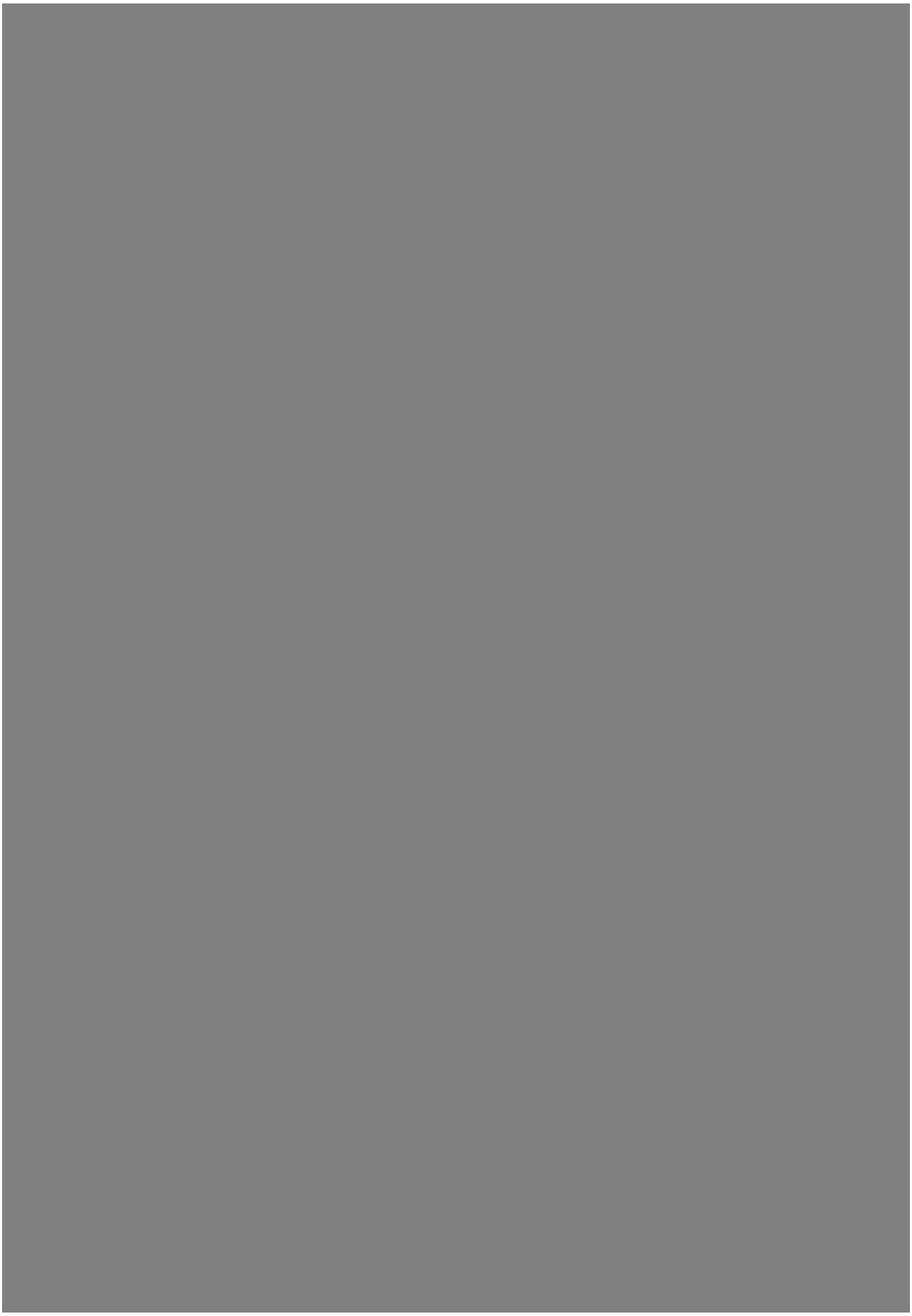


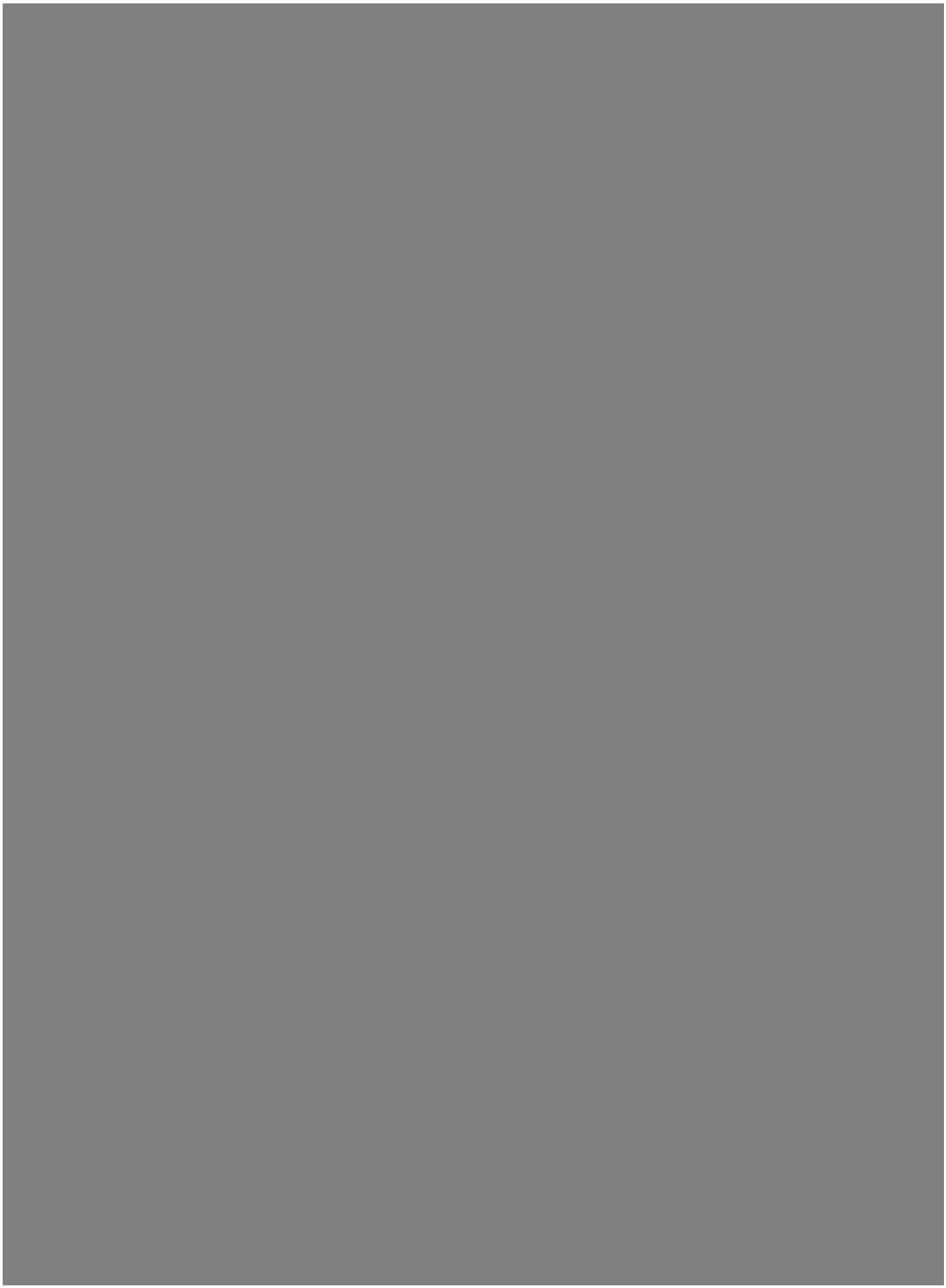












12.0 Appendix 3 : Electronic search for Medline for CKD outcomes

Count	Searches	Results
1	exp Renal Insufficiency Chronic/	84840
2	chronic kidney disease.mp.	18247
3	chronic kidney disease$.mp.	18745
4	exp Kidney Failure, Chronic/	78448
5	chronic kidney failure.mp.	975
6	chronic kidney failure$.mp.	976
7	chronic renal failure.mp.	20307
8	chronic renal failure$.mp.	20319
9	end stage kidney disease.mp.	997
10	end stage kidney disease$.mp.	1005
11	esrd.mp.	9767
12	esrd$.mp.	9769
13	chronic kidney insufficiency.mp.	195
14	chronic kidney insufficiency$.mp.	195
15	end stage renal disease.mp.	19454
16	end stage renal disease$.mp.	19625
17	end stage renal failure.mp.	4888
18	end stage renal failure$.mp.	4891
19	kidney failure.mp.	81179
20	kidney failure$.mp.	81828
21	renal insufficiency/	10919
22	renal failure.mp.	71184
23	renal failure$.mp.	71271
24	kidney insufficiency.mp.	577
25	kidney insufficiency$.mp.	577
26	exp renal dialysis/	93204
27	renal dialysis.mp.	74276
28	renal dialysis$.mp.	74276
29	extracorporeal dialysis.mp.	161
30	extracorporeal dialysis$.mp.	162
31	hemodialysis.mp.	47033
32	hemodialysis$.mp.	47039
33	haemodialysis.mp.	11565
34	haemodialysis$.mp.	11576
35	exp peritoneal dialysis/	22761
36	peritoneal dialysis.mp.	26208
37	peritoneal dialysis$.mp.	26210
38	renal disease.mp.	39428
39	renal disease$.mp.	43692
40	exp kidney diseases/	414228
41	kidney disease.mp.	30990
42	kidney disease$.mp.	103750
43	nephropathy.mp.	37719

Count	Searches	Results
44	nephropathy$.mp.	37721
45	exp diabetic nephropathies/	19707
46	diabetic nephropathy.mp.	11909
47	diabetic nephropathy$.mp.	11909
48	exp kidney transplantation/	79900
49	kidney transplantation.mp.	81478
50	kidney transplantation$.mp.	81539
51	renal transplant.mp.	18888
52	renal transplant$.mp.	36596
53	exp dialysis/	22124
54	dialysis.mp.	132315
55	dialysis$.mp.	132329
56	exp renal insufficiency/	127094
57	renal insufficiency.mp.	33560
58	renal insufficiency$.mp.	33571
59	EGFR.mp.	27296
60	EGFR$.mp.	27669
61	exp glomerular filtration rate/	33096
62	glomerular filtration rate.mp.	42765
63	glomerular filtration rate$.mp.	43111
64	exp creatinine/	46991
65	creatinine.mp.	95418
66	creatinine$.mp.	95673
67	serum creatinine.mp.	26670
68	serum creatinine$.mp.	26781
69	serum creatinine clearance.mp.	53
70	serum creatinine clearance$.mp.	58
71	exp albuminuria/	12045
72	albuminuria$.mp.	14614
73	exp proteinuria/	32445
74	proteinuria.mp.	36905
75	proteinuria$.mp.	36921
76	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75	683333
77	exp diabetes mellitus, type 2/	87630
78	diabetes mellitus type 2.mp.	87782
79	type 2 diabetes.mp.	60366
80	type 2 diabetes$.mp.	60434
81	niddm.mp.	6673
82	niddm$.mp.	6704
83	exp diabetes insipidus/	7011

Count	Searches	Results
84	diabetes insipidus.mp.	8589
85	diabetes insipidus$.mp.	8589
86	77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85	113193
87	exp glucose intolerance/	6360
88	impaired glucose tolerance.mp.	8065
89	impaired glucose tolerance$.mp.	8066
90	glucose intolerance.mp.	11300
91	glucose intolerance$.mp.	11302
92	exp prediabetic state/	3852
93	prediabetes.mp.	1633
94	prediabetic state.mp.	4004
95	prediabetic state$.mp.	4053
96	exp blood glucose/	130285
97	blood glucose.mp.	147811
98	blood glucose$.mp.	147841
99	glucose metabolism.mp.	23463
100	glucose metabolism$.mp.	23713
101	exp glucose tolerance test/	29397
102	glucose tolerance test.mp.	33799
103	glucose tolerance test$.mp.	34976
104	OGTT.mp.	5382
105	OGTT$.mp.	5481
106	exp Hyperglycemia/	26636
107	hyperglycemia.mp.	37642
108	hyperglycemia$.mp.	37682
109	hyperglycaemia.mp.	6827
110	hyperglycaemia$.mp.	6840
111	impaired fasting glucose.mp.	2336
112	impaired fasting glucose$.mp.	2336
113	postprandial hyperglycemia.mp.	911
114	postprandial hyperglycaemia.mp.	283
115	exp hemoglobin a, glycosylated/	23254
116	hemoglobin a, glycosylated.mp.	23255
117	hemoglobin a, glycosylated$.mp.	23255
118	Haemoglobin a, glycosylated.mp.	1
119	HbA1c.mp.	13678
120	HbA1c$.mp.	13741
121	glycemic abnormality.mp.	7
122	Glycaemic abnormality.mp.	0
123	Fasting plasma glucose.mp.	7250
124	Fasting plasma glucose$.mp.	7259
125	87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102 or 103 or 104 or 105 or 106 or 107 or 108 or 109 or 110 or 111 or 112 or 113 or 114 or 115 or 116 or 117 or 118 or 119 or 120 or 121 or 122 or 123 or 124	223393
126	76 and 86 and 125	4367

Count	Searches	Results
127	exp cohort studies/	1387399
128	cohort$.tw.	263907
129	controlled clinical trial.pt.	88411
130	epidemiologic methods/	29569
131	exp case-control studies/	688110
132	(case$ and control$).tw.	311313
133	127 or 128 or 129 or 130 or 131 or 132	1909067
134	cohort studies/	170386
135	longitudinal studies/	87299
136	follow-up studies/	504746
137	prospective studies/	376036
138	retrospective studies/	512412
139	cohort.ti,ab.	239106
140	longitudinal.ti,ab.	132074
141	prospective.ti,ab.	339467
142	retrospective.ti,ab.	264289
143	Case-Control Studies/	188791
144	Control Groups/	1435
145	Matched-Pair Analysis/	4154
146	retrospective studies/	512412
147	((case* adj5 control*) or (case adj3 comparison*) or control group*).ti,ab.	374925
148	127 or 128 or 129 or 130 or 131 or 132 or 133 or 134 or 135 or 136 or 137 or 138 or 139 or 140 or 141 or 142 or 143 or 144 or 145 or 146 or 147	2266708
149	126 and 148	1563
150	Remove duplicates from 149	1560

13.0 Appendix 4: Sensitive search of Medline for CKD outcomes

Count	Searches	Results
1	exp Renal Insufficiency Chronic/	98188
2	chronic kidney disease.mp.	29057
3	chronic kidney disease$.mp.	29695
4	exp Kidney Failure, Chronic/	85933
5	chronic kidney failure.mp.	1102
6	chronic kidney failure$.mp.	1103
7	chronic renal failure.mp.	21495
8	chronic renal failure$.mp.	21510
9	end stage kidney disease.mp.	1526
10	end stage kidney disease$.mp.	1536
11	esrd.mp.	11723
12	esrd$.mp.	11727
13	chronic kidney insufficiency.mp.	197
14	chronic kidney insufficiency$.mp.	197
15	end stage renal disease.mp.	23179
16	end stage renal disease$.mp.	23319
17	end stage renal failure.mp.	5301
18	end stage renal failure$.mp.	5304
19	kidney failure.mp.	89943
20	kidney failure$.mp.	89958
21	renal insufficiency/	14052
22	renal failure.mp.	77621
23	renal failure$.mp.	77723
24	kidney insufficiency.mp.	609
25	kidney insufficiency$.mp.	609
26	exp renal dialysis/	103600
27	renal dialysis.mp.	83217
28	renal dialysis$.mp.	83217
29	extracorporeal dialysis.mp.	173
30	extracorporeal dialysis$.mp.	174
31	hemodialysis.mp.	52361
32	hemodialysis$.mp.	52366
33	haemodialysis.mp.	12671
34	haemodialysis$.mp.	12682
35	exp peritoneal dialysis/	24546
36	peritoneal dialysis.mp.	28321
37	peritoneal dialysis$.mp.	28326
38	renal disease.mp.	45665
39	renal disease$.mp.	50430
40	exp kidney diseases/	463998
41	kidney disease.mp.	44884
42	kidney disease$.mp.	125612
43	nephropathy.mp.	43133

Count	Searches	Results
44	nephropathy$.mp.	43138
45	exp diabetic nephropathies/	22277
46	diabetic nephropathy.mp.	13892
47	diabetic nephropathy$.mp.	13893
48	exp kidney transplantation/	86693
49	kidney transplantation.mp.	88628
50	kidney transplantation$.mp.	88690
51	renal transplant.mp.	20699
52	renal transplant$.mp.	39656
53	exp dialysis/	22949
54	dialysis.mp.	144958
55	dialysis$.mp.	144967
56	exp renal insufficiency/	147955
57	renal insufficiency.mp.	43607
58	renal insufficiency$.mp.	43608
59	EGFR.mp.	37069
60	EGFR$.mp.	37583
61	exp glomerular filtration rate/	37849
62	glomerular filtration rate.mp.	49606
63	glomerular filtration rate$.mp.	50041
64	exp creatinine/	51891
65	creatinine.mp.	108189
66	creatinine$.mp.	108453
67	serum creatinine.mp.	31076
68	serum creatinine$.mp.	31197
69	serum creatinine clearance.mp.	59
70	serum creatinine clearance$.mp.	64
71	exp albuminuria/	13604
72	albuminuria$.mp.	16772
73	exp proteinuria/	36295
74	proteinuria.mp.	41047
75	proteinuria$.mp.	41065
76	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75	765478
77	exp glucose intolerance/	7404
78	impaired glucose tolerance.mp.	9148
79	impaired glucose tolerance$.mp.	9149
80	glucose intolerance.mp.	12943
81	glucose intolerance$.mp.	12945
82	exp prediabetic state/	5110
83	prediabetes.mp.	2524
84	prediabetic state.mp.	5243

Count	Searches	Results
85	prediabetic state$.mp.	5279
86	exp blood glucose/	147244
87	blood glucose.mp.	168289
88	blood glucose$.mp.	168320
89	glucose metabolism.mp.	27693
90	glucose metabolism$.mp.	27738
91	exp glucose tolerance test/	32128
92	glucose tolerance test.mp.	37296
93	glucose tolerance test$.mp.	38628
94	OGTT.mp.	6360
95	OGTT$.mp.	6480
96	exp Hyperglycemia/	31346
97	hyperglycemia.mp.	44272
98	hyperglycemia$.mp.	44321
99	hyperglycaemia.mp.	7873
100	hyperglycaemia$.mp.	7888
101	impaired fasting glucose.mp.	2874
102	impaired fasting glucose$.mp.	2875
103	postprandial hyperglycemia.mp.	1080
104	postprandial hyperglycaemia.mp.	318
105	exp hemoglobin a, glycosylated/	28871
106	hemoglobin a, glycosylated.mp.	28873
107	hemoglobin a, glycosylated$.mp.	28873
108	Haemoglobin a, glycosylated.mp.	1
109	HbA1c.mp.	19317
110	HbA1c$.mp.	19411
111	glycemic abnormality.mp.	9
112	Glycaemic abnormality.mp.	2
113	Fasting plasma glucose.mp.	8981
114	Fasting plasma glucose$.mp.	8992
115	77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102 or 103 or 104 or 105 or 106 or 107 or 108 or 109 or 110 or 111 or 112 or 113 or 114	257303
116	exp cohort studies/	1652390
117	cohort$.tw.	348084
118	controlled clinical trial.pt.	93340
119	epidemiologic methods/	30928
120	exp case-control studies/	850703
121	(case$ and control$).tw.	363820
122	cohort studies/	208646
123	longitudinal studies/	107497
124	follow-up studies/	577197
125	prospective studies/	449579
126	retrospective studies/	634617
127	cohort.ti,ab.	315613

Count	Searches	Results
128	longitudinal.ti,ab.	160875
129	prospective.ti,ab.	407292
130	retrospective.ti,ab.	331108
131	Case-Control Studies/	232869
132	Control Groups/	1585
133	Matched-Pair Analysis/	4559
134	retrospective studies/	634617
135	((case* adj5 control*) or (case adj3 comparison*) or control group*).ti,ab.	448137
136	116 or 117 or 118 or 119 or 120 or 121 or 122 or 123 or 124 or 125 or 126 or 127 or 128 or 129 or 130 or 131 or 132 or 133 or 134 or 135	2686063
137	76 and 115 and 136	6187
138	limit 137 to humans	5737
139	limit 138 to yr="1946 - 2015"	5487
140	limit 139 to "adult (19 to 44 years)"	2984

14.0 Appendix 5: Review eligibility criteria checklist

Study design	Cohort studies Case-control studies
Study characteristics	Full articles Conference proceedings Grey literature Theses/Dissertations Other (please specify)
Participants	Studies where some participants are aged (18 to 40 years) With IGR With pre-diabetes (can refer to either IGT or Impaired Fasting Glucose (IFG). With metabolic syndrome (where IGR is part of metabolic syndrome) Free from CKD at baseline
Comparator	Participant with normoglycaemia Participants with diabetes
Outcome	Chronic kidney disease [eGFR stages: 3A, 3B, 4 and 5] Albuminuria Proteinuria ≥1 ACR (albumin creatinine ratio ≥30mg/mmol PCR (protein creatinine ratio ≥50mg/mmol SCr (serum creatinine) data CrCl (creatinine clearance) data

15.0 Appendix 6: Data extraction form

Abbreviation

IGR	Impaired Glucose Regulation
CKD	Chronic Kidney Disease
T2DM	Type 2 Diabetes
SCr	Serum Creatinine
CrCl	Creatinine Clearance
ESRD	End Stage Renal Disease
ACR	Albumin Creatinine Ratio
PCR	Protein Creatinine Ratio
EX	Excluded
NR	Not Reported
eGFR	Estimated Glomerular Filtration Rate

Eligibility criteria for title and abstract screening phase

Study design	**Assessment**	**Comment**
Is it: [1] A cohort study (Prospective or Retrospective) [2] A case-control or nested case-control study	Yes No Unclear	
Population		
[1] Were patients diagnosed with IGR? [2] Were patients followed up for T2DM NB: Please answer Yes if IGR diagnosed in a sub group	Yes No Unclear	
Are patients aged (18-40) years NB: Please answer Yes if mixed age population	Yes No Unclear	
Outcomes		
Did the study report any of the following outcome:	Yes No	

<table>
<tr><th>Study design</th><th>Assessment</th><th>Comment</th></tr>
<tr><td>
<table>
<tr><td>Stage</td><td>eGFR (ml/min/1.73m^2)</td></tr>
<tr><td>3A</td><td>45-59</td></tr>
<tr><td>3B</td><td>30-44</td></tr>
<tr><td>4</td><td>15-29</td></tr>
<tr><td>5</td><td><15</td></tr>
</table>
ESRD (stage 5)
ACR - >30mg/mmol
PCR - >45mg/mmol
SCr measures
CrCl measures
</td><td>Unclear</td><td></td></tr>
<tr><td>Follow up</td><td></td><td></td></tr>
<tr><td>Were the patients followed up and adequate measures taken

NB: Please answer Yes if adequate measure were taken and key characteristics described</td><td>Yes
No
Unclear</td><td></td></tr>
<tr><td>Final decision (please tick)</td><td>Include
Exclude
Unclear</td><td></td></tr>
</table>

Exclusion criteria

Reasons for exclusion of study from review (please circle where appropriate)	
Methods	Not a cohort/case-control study
Patients	No IGR/no T2DM follow up/T2DM or Type 1 diabetes diagnosis /CKD /wrong age group
Outcomes	No relevant outcomes assessed No data for relevant subgroup extractable
Follow up period	No follow up
Other	Duplicate publication Other

Inclusion criteria

Specific inclusion criteria (please include if answer is Yes to all questions below)	
Eligibility criteria	
Satisfaction of eligibility criteria	Yes No Unclear
Effect sizes	
Is there sufficient reporting of statistics or data to calculate effect sizes	Yes No Unclear

Organisation

Organisational aspect		**Exclude**		**Include**
Reviewer/date:		**Checked by:**		
Author/Year				
Journal/Source				
Country of origin				
Publication type	Full text/Abstract/Book chapter/progress report/ Other – please specify			
Fate	Decision pending/Check references/Use for discussion/EX without listing/EX with listing Other – please specify			
Notes				

Study characteristics

<table>
<tr><th colspan="2">General study characteristics (please circle where appropriate)</th></tr>
<tr><td>Location of study</td><td></td></tr>
<tr><td>Study aims</td><td>Reported/NR</td></tr>
<tr><td>Dates of recruitment</td><td>From _____ to _____
Median (range): #
Mean:#</td></tr>
<tr><td>Length of follow up of CKD outcome + length of follow up of study</td><td>From _____ to _____
Median (range): #
Mean:#</td></tr>
<tr><td>Outcomes assessed</td><td>Did the study report any of the following outcome:<table><tr><td>Stage</td><td>eGFR (ml/min/1.73m^2)</td></tr><tr><td>3A</td><td>45-59</td></tr><tr><td>3B</td><td>30-44</td></tr><tr><td>4</td><td>15-29</td></tr><tr><td>5</td><td><15</td></tr></table>End Stage Renal Disease (stage 5)
Albumin creatinine ratio - >30mg/mmol
Protein creatinine ratio - >45mg/mmol
Serum creatinine measures
Creatinine clearance measures
Other (please specify):</td></tr>
<tr><td>Outcome definition</td><td></td></tr>
<tr><td>Relationship between outcome and relevant factor</td><td>Is the relationship statistically significant?
Yes/No
RR/OR: #

If No, is it due to
Low powered or inconclusive study/A true negative study</td></tr>
<tr><td>Power calculation</td><td>Yes/No/Not reported

Calculated sample size: #
Sample size achieved: Yes/No</td></tr>
<tr><td></td><td></td></tr>
<tr><td>Funding</td><td>Unclear
NR
Please state where reported</td></tr>
<tr><td>Conflict of interest statement</td><td>Yes/No/NR</td></tr>
</table>

<table>
<tr><th colspan="4">Baseline characteristics of patients (please circle where appropriate)</th></tr>
<tr><td></td><td>Exposure</td><td>Control</td><td>Notes: any relationship with outcomes? Yes/No/NR
If Yes please state if statistically significant and RR/OR values</td></tr>
<tr><td colspan="4">Overall comment: Significant/In significant</td></tr>
<tr><td>Number of patients</td><td></td><td></td><td></td></tr>
<tr><td>Age range (if reported)
Mean</td><td></td><td></td><td></td></tr>
<tr><td>Ethnicity
No%</td><td></td><td></td><td></td></tr>
<tr><td>Gender
No%</td><td>Male:
Female:</td><td>Male:
Female:</td><td></td></tr>
<tr><td>No of patients screened for IGR</td><td></td><td></td><td></td></tr>
<tr><td>No of patients recruited</td><td></td><td></td><td></td></tr>
<tr><td>No of patients allocated</td><td></td><td></td><td></td></tr>
<tr><td>No of patients evaluated</td><td></td><td></td><td></td></tr>
<tr><td>No of dropouts</td><td></td><td></td><td></td></tr>
<tr><td>Reasons for dropouts</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>Number of protocol violations</td><td></td><td></td><td></td></tr>
<tr><td></td><td colspan="3"></td></tr>
<tr><td>Definition of IGR

<table><tr><td>Fasting plasma glucose</td><td>6.1-6.9 mmol/L</td></tr><tr><td>Oral Glucose Tolerance test (2h value)</td><td>7.8-11.0 mmol/L</td></tr><tr><td>HbA1c</td><td>42-47 mmol/mol</td></tr></table>
Please circle all that applies and list all</td><td></td><td></td><td></td></tr>
<tr><td>Status of patient at recruitment

Any treatment for any comorbidities

If treated:
Please state

What treatment

Duration</td><td></td><td></td><td></td></tr>
<tr><td>Adverse event?
Yes/No

If Yes please state</td><td></td><td></td><td></td></tr>
</table>

Observational study characteristics (please circle where appropriate)	
Sample size	
Number of excluded patients	
Recruitment method	
Type of observational study	Cohort studies (retrospective/prospective) Case-control studies/nested case-control
Are group comparable?	Yes/No If No, please specify
Any confounders?	Yes/No If No, please specify
Analysis	
Drop outs stated	Yes/No If Yes: # in each group

Outcome details

The following table have to be copied for every relevant outcome assessed *(please fill out fields only where applicable)*

Outcomes assessed (please state where relevant)	
Definition of each outcome	
Time of assessment of each outcome (post IGR)	
Timing of assessment	
Length of follow up for each outcome	
Method of measurement	
No of patients evaluated for each outcome, as stated above	

Methodological quality summary for observational studies					
Reviewer/Date:			Checked by:		
Contents (please refer to tables below for guidance	Yes	Partly	No	Unsure	Comments
Study participation					
Study attrition					
Measurement of prognostic factors					
Measurement and controlling for confounding variables					
Measurement of outcomes					
Analysis approach					
Summarised validity	Low risk of bias		Moderate risk of bias		High risk of bias
Remarks:					

16.0 Appendix 7: Quality assessment form adapted from Ottawa-Newcastle scale (NOS) for assessing non-randomised studies

<table>
<tr><th></th><th></th><th>Yes/No/Unclear</th></tr>
<tr><td>Selection of participants</td><td>[1] Was the inclusion/exclusion clearly described?
(for example, age, diagnosis status, IGT/IFG)

[2] Was inclusion/exclusion assessed using valid and reliable measures? (For example, if there are important inclusion/exclusion criteria that are not directly related to exposure and outcome and for which the accuracy of measurement may need scrutiny, e.g. age, diagnosis).

[3] Was recruitment strategy clearly described?

[4] Did the investigators ensure that the exposed/unexposed group were comparable (for example, did they use stratification, matching or propensity Score).</td><td></td></tr>
<tr><td>Adequate description of study population</td><td>[1] Was study population well characterised?
• Age
• Sex
• Ethnicity
• Suitable definition of IGT/IFG</td><td></td></tr>
<tr><td>Validated method for ascertaining exposure</td><td>[1] Was the method used to ascertain exposure clearly defined?
[2] Was a valid and reliable measure used to ascertain exposure?
(For example what diagnostic test was used to confirm IGT/IFG)
<table><tr><td>Fasting Plasma Glucose</td><td>6.1 – 6.9 mmol/L</td></tr><tr><td>Oral Glucose Tolerance Test (2h value)</td><td>7.8 – 11.0 mmol/L</td></tr><tr><td>HbA1c</td><td>42 – 47 mmol/mol / 5.4-6.7%</td></tr></table></td><td></td></tr>
<tr><td>Validated method to confirm outcome</td><td>[1]Was valid and reliable measures used to ascertain outcome? For example
<table><tr><td>Stage</td><td>eGFR (ml/min/1.73m^2)</td></tr><tr><td>1</td><td>≥90</td></tr><tr><td>2</td><td>60-89</td></tr><tr><td>3A</td><td>45-59</td></tr><tr><td>3B</td><td>30-44</td></tr><tr><td>4</td><td>15-29</td></tr><tr><td>5</td><td><15</td></tr></table>ACR - >30mg/mmol
PCR - >45mg/mmol
SCr measures
CrCl measures</td><td></td></tr>
<tr><td>Adequate follow up period</td><td>[1] Was follow up long enough for the outcome to occur?

[2] Was the follow up period the same across all groups?

[3] Were differences in follow-up adjusted for using statistical techniques, e.g., survival analysis?</td><td></td></tr>
<tr><td>Completeness of follow-up (Attrition)</td><td>[1] Were drop-out rates and reasons for drop-out similar across exposed and unexposed?

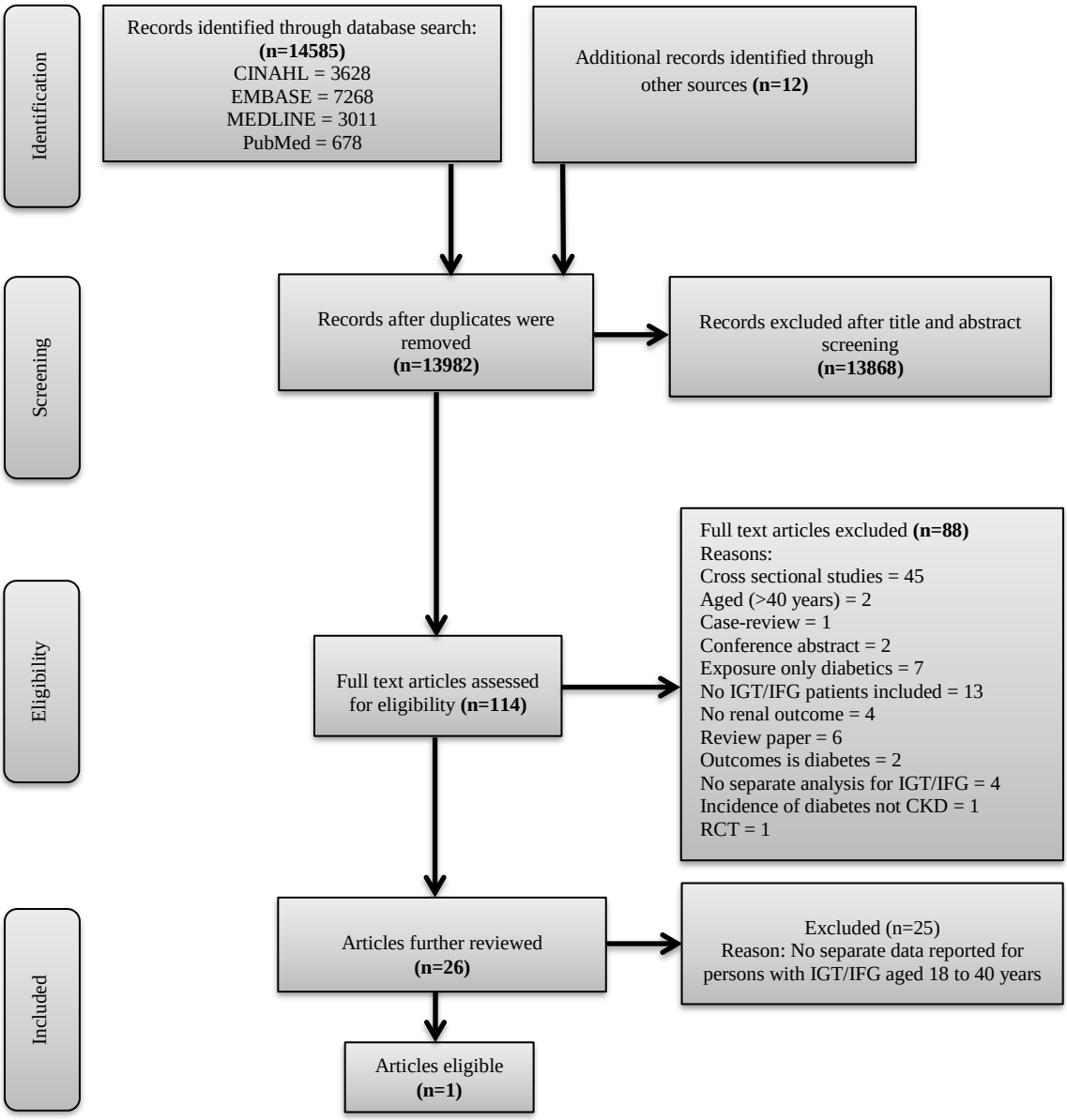
[2] Were numbers of dropouts/withdrawals documented at each time point?</td><td></td></tr>
</table>

		Yes/No/Unclear
Analysis controls for confounding	[1] Does the study identify and control for important confounding variables and effect modifiers?	
Sample size calculated	[1] Is the sample size adequate? [2] Did the study describe how the sample size was calculated? • Did the investigators conduct a power analysis to determine the adequacy of study group sizes for the outcome of interest? • Was the sample size large enough to detect differences in event or a significant OR/RR between groups? Mean (+/-SE) change in GFR	
Analytical methods appropriate	[1] Was the kind of analysis done appropriate for the kind of outcome data? For example, • Dichotomous – logistic regression, survival analysis • Categorical – mixed model for categorical outcomes • Continuous – Mixed model, ANCOVA • Mean change (+/-SE) [2] Was loss to follow up accounted for in the analysis (For example, through sensitivity analysis)	

Overall appraisal: Include ☐ Exclude ☐ Seek further info ☐

Comments (including reasons for exclusion):

17.0 Appendix 8: Sensitivity search PRISMA flow diagram



18.0 Appendix 9: Characteristics of studies reporting IGT/IFG compared to normoglycaemia and development of CKD

Author/Year (Reference)	Country	Study type	Mean age (+/- SD) or range (years) or percentages	Follow-up	Source	Definition of CKD	No. of individuals with IGT/IFG	Definition of exposure	Measures of association (OR/RR/HR/IRR) (95% CI)	Adjustments
Nelson et al (1996)	US	Prospective cohort	aged (18 to 60 years)	4 years	Community based population	ACR	29/60	IGT	Mean (+/-SE) change in GFR	N/A
Fox et al (2005)	US	Prospective cohort	aged (28 to 62 years)	7 years	Framingham Offspring study	eGFR<60ml/min/1.73m^2	704/3,102	IGT/IFG	Unadjusted OR: 1.65(1.16 to 2.36) Adjusted OR: 0.98(0.67 to 1.45)	Age, sex, baseline GFR, SBP, hypertension treatment, smoking, BMI, total & HDL cholesterol, MI, congestive heart failure
Meigs et al (2002)	US	Prospective cohort	Aged (≥34 years)	24 years	Framingham Offspring study	ACR	2,501/5,330	IGT/IFG	Adjusted OR: Men – 1.06(1.03 to 1.09) Women – 1.07 (1.04 to 1.10)	Age, SBP, BMI, smoking, ACE inhibitor, total cholesterol, HDL, triglyceride, hypertensive drugs.
Nelson et al (1999)	US	Prospective cohort	Aged (18 to 60 years)	4 years	Community based population	ACR	27/80	IGT	Mean (+/-SE) change in GFR	N/A
Nelson et al (1989)	US	Prospective cohort	Aged (≥15 years)	6 years	Community based population	ACR	217/2,945	IGT	Adjusted OR: 1.5(1.0 to 2.2)	Age, sex, BP
Yokoyama et al (2009)	Japan	Prospective cohort	Aged 61 (+/-12)	Median 3 years	Outpatient clinic	eGFR<60ml/min/1.73m^2	194/1,117	IGT/IFG	Unadjusted OR: 1.31(0.91 to 1.87)	No adjustment
Tozawa et al (2007)	Japan	Prospective cohort	Aged (19 to 84 years)	5 years	Community based population	GFR<60ml/min/1.73m^2	884/7,255	IGT/IFG	Unadjusted RR: 1.28(0.94 to 1.75) Adjusted RR: 1.22(0.89 to 1.07)	Age, sex, current cigarette smoking, alcohol drinking habit
Nelson et al (1993)	US	Prospective cohort	Aged range (15 to 80)	Mean (5.2 years)	Community based population	ACR	412/902	IGT	Unadjusted OR: 0.8(0.4 to 1.5)	No adjustment
Rashidi et al (2007)	Iran	Prospective cohort	Aged 48 (+/-12)	3 years	Community based population	CrCL	1,010/5,617	IFG	Unadjusted OR: 1.02(0.99 to 1.04)	No adjustment
Kitiyakara et al (2007)	Thailand	Prospective cohort	Aged 43.0 (+/-4.8)	12 years	Targeted population (Electricity generating authority of Thailand	eGFR<60ml/min/1.73m^2	227/2,294	IGT/IFG	Unadjusted OR based on IDF definition – 1.59(0.99 to 2.54)Unadjusted OR based on NCEP definition – 2.44(1.29 to 4.60)Adjusted OR based on IDF definition – 1.38(0.85 to 2.22)Adjusted OR based on NCEP definition – 1.97(1.03 to 3.78)	Age, sex and smoking status
Sun et al (2010)	Taiwan	Retrospective cohort	Aged (20 to 74 years)	3.7 (+/-2.4)	Population database	eGFR<60ml/min/1.73m^2	17,039/135,963	IGT/IFG	Unadjusted HR: 1.26(1.22 to 1.32)Adjusted HR: 0.97(0.93 to 1.01)	Age, sex, check-up centres and current smoking
Yang et al (2012)	Taiwan	Prospective cohort	Aged 50.51 (+/-13.98)	Median (5.4 years)	Population survey	eGFR<60ml/min/1.73m^2	637/4,885	IGT/IFG	Unadjusted HR: 2.47(1.83 to 3.32) Adjusted HR: 1.33(1.03 to 1.89)	Age, sex, BMI, serum level, total cholesterol, BP, triglyceride, HDL
Kovacs et al (2013)	Hungary	Prospective cohort	36.4 (+/-13.0)	Month: 146(+/-112.5)	Targeted population Nephrology dept. University of Pecs	eGFR<60ml/min/1.73m^2 SCr	107/330	IGT	Unadjusted HR: 0.67(0.41 to1.10) -SCr HR: 0.60(0.42 to 0.86) - eGFR<60ml/min/1.73m^2 HR: 0.67(0.42 to 1.09) - eGFR<30ml/min/1.73m^2 HR: 0.88(0.49 to 1.60) - eGFR<15ml/min/1.73m	No adjustments

Author/Year (Reference)	Country	Study type	Mean age (+/- SD) or range (years) or percentages	Follow-up	Source	Definition of CKD	No. of individuals with IGT/IFG	Definition of exposure	Measures of association (OR/RR/HR/IRR) (95% CI)	Adjustments
Watanabe et al (2010)	Japan	Prospective cohort	50.0 (+/-7.8)	5.8 (+/-2.4) years	Community based population	eGFR<60ml/min/1.73m^2	1,082/15,971	IGT	Adjusted HR: 1.94(1.06 to 3.54) - eGFR<60ml/min/1.73m^2	Sex and age
Ryu et al (2009)	South Korea	Prospective cohort	37.4 (+/-4.8)	3.8 years	Comprehensive health check	eGFR<60ml/min/1.73m^2	787/10,685	IGT/IFG	Adjusted HR: 1.04(0.60 to 1.79)	Age, baseline GFR, glutamyltranspeptide, uric acid, triglyceride, HDL cholesterol, BP, obesity
Jee et al (2005)	South Korea	Prospective cohort	Aged (35 to 59 years)	10 years	Hospital setting	Dipstick proteinuria (≥1+ and 2+)	3,913/57,318	IGT/IFG	Adjusted RR: 1.7(1.4 to 1.9) – MenAdjusted RR: 1.5(1.0 to 2.2) - Women	BMI, total cholesterol, SBP
Tohidi et al (2012)	Iran	Prospective cohort	Aged (≥20 years)	10 years	Community based population	eGFR<60ml/min/1.73m^2	3,313/8,395	IGT/IFG	Unadjusted OR: 0.95(0.73 to 1.24)	No adjustments
Schottker et al (2013)	Germany	Prospective cohort	Aged (50 – 74 years)	8 years	ESTHER study	eGFR<60ml/min/1.73m^2	258/3,538	IGT/IFG	Unadjusted RR: 1.33(1.03 to 1.73) Adjusted RR: 1.06(0.71 to 1.32)	BMI, BP, cholesterol, antihypertensive drug, statins, smoking, history of CVD
Bonnet et al (2006)	France	Prospective cohort	Aged (30-64 years)	6 years	DESIR study	Urinary albumin 20mg/L Dipstick positive proteinuria	760/2,738	IFG	Adjusted OR: 1.87(1.25 to 2.81) – Men Adjusted OR: 1.40(0.80 to 2.46) - Women	Age, ACE inhibitors, smoking, fibrinogen level
Lucove et al (2008)	USA	Prospective cohort	Aged (45-74 years)	9 years	Strong heart study	ACR: ≥30mg/g eGFR<60ml/min/1.73m^2	896/2,380	IFG/IGT	Adjusted HR: 1.3(1.1 to 1.6)	age, sex, study centre, education, and smoking
Nand et al (2015)	India	Case-control	Aged (≥20 years)	N/A	Hospital outpatient	eGFR<60ml.min/1.73m^2	150/300	IFG	Unadjusted OR: 2.91(1.80 to 4.80)	No adjustment
Halbesma et al (2008)	Netherlands	Prospective cohort	49 (+/-12)	6.5 years	PREVEND study	eGFR<60ml.min/1.73m^2	220/5,488	IGT/IFG	Change in GFR overtime	N/A
Ninomiya et al (2006)	Japan	Prospective cohort	Aged (≥40 years)	5 years	Community based population	eGFR<60ml.min/1.73m^2	353/1,440	IFG/IGT	Adjusted OR: 2.08(1.23 to 3.52)	Age, sex, hyperinsulinemia, GFR, proteinuria, serum albumin, cholesterol, haemoglobin, alcohol, tobacco use
Carson et al (2015)	USA	Prospective cohort	Aged (18-30 years)	25 years	CARDIA study	eGFR<60ml.min/1.73m^2	2,174/5,115	IGT/IFG	Unadjusted RR: 1.59(1.04 to 2.43) Adjusted RR: 1.11(0.71 to 1.72)	Age, race, sex, education

19.0 Appendix 10: Characteristics of studies reporting IGT/IFG compared to T2DM and development of CKD

Author/Year (Reference)	Country	Study type	Mean age (+/- SD) or range (years) or percentages	Follow-up	Source	Definition of CKD	No. of individuals with IGT/IFG	Definition of exposure	Measures of association (OR/RR/HR/IRR) (95% CI)	Adjustments
Kim et al (2010)	US	Prospective cohort	Aged (5 to 19 years)	25.2 years	Community based population	ACR	2,534/5,200	IGT	Incidence (cases/1000 PY) 0.13% (IGT) 2.4% (T2DM)	Age and sex
Iseki et al (2004)	Japan	Prospective cohort	Aged (>20 years)	7.75 years	Community based population	SCr	74,440/152,969	IGT/IFG	Adjusted OR: 3.01(1.74 to 5.52)	SBP, DBP, BMI, total cholesterol, triglyceride, serum creatinine

[5] OR, odds ratio; IRR, incidence rate ratio; HR hazard ratio; GFR, glomerular filtration rate; ACR, albumin creatinine ratio; SCr, serum creatinine; CrCl, creatinine clearance; SBP, systolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; MI, myocardial infarction; HDL, high density lipoprotein; BP, blood pressure; DBP, diastolic blood pressure; SD, standard deviation; SE, standard error

20.0 Appendix 11: Read code description

Read codes used to identify cases of IGR

Read code	Description
44V2.00	Glucose tolerance test impaired
C11y200	Impaired glucose tolerance
C11y300	Impaired fasting glycaemia
C11y400	Impaired glucose regulation
C11y500	Pre-diabetes
6AC..00	Review of impaired glucose tolerance
9NS0400	Referral for impaired glucose tolerance management offered
R102.00	[D]Glucose tolerance test abnormal
R102.11	[D]Prediabetes
R102.12	[D]Impaired glucose tolerance test
R10D000	[D]Impaired fasting glycaemia
R10D011	[D]Impaired fasting glucose
R10E.00	[D]Impaired glucose tolerance
C313500	Glucose intolerance

Abbreviation: [D], diagnosis

Read code used to identify cases of T2DM

Note: This list include codes which do not specifically specify diabetes type

Read Code	Description
C10..00	Diabetes mellitus
C109J00	Insulin treated Type 2 diabetes mellitus
C109K00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus
C10C.00	Diabetes mellitus autosomal dominant
C10D.00	Diabetes mellitus autosomal dominant type 2
C10F.00	Type 2 diabetes mellitus
C10F.11	Type II diabetes mellitus
C10F000	Type 2 diabetes mellitus with renal complications
C10F011	Type II diabetes mellitus with renal complications
C10F100	Type 2 diabetes mellitus with ophthalmic complications
C10F111	Type II diabetes mellitus with ophthalmic complications
C10F200	Type 2 diabetes mellitus with neurological complications
C10F211	Type II diabetes mellitus with neurological complications
C10F300	Type 2 diabetes mellitus with multiple complications
C10F311	Type II diabetes mellitus with multiple complications
C10F400	Type 2 diabetes mellitus with ulcer
C10F411	Type II diabetes mellitus with ulcer
C10F500	Type 2 diabetes mellitus with gangrene
C10F511	Type II diabetes mellitus with gangrene
C10F600	Type 2 diabetes mellitus with retinopathy
C10F611	Type II diabetes mellitus with retinopathy
C10F700	Type 2 diabetes mellitus - poor control
C10F711	Type II diabetes mellitus - poor control
C10F900	Type 2 diabetes mellitus without complication
C10F911	Type II diabetes mellitus without complication
C10FA00	Type 2 diabetes mellitus with mononeuropathy
C10FA11	Type II diabetes mellitus with mononeuropathy
C10FB00	Type 2 diabetes mellitus with polyneuropathy
C10FB11	Type II diabetes mellitus with polyneuropathy
C10FC00	Type 2 diabetes mellitus with nephropathy
C10FC11	Type II diabetes mellitus with nephropathy
C10FD00	Type 2 diabetes mellitus with hypoglycaemic coma
C10FD11	Type II diabetes mellitus with hypoglycaemic coma
C10FE00	Type 2 diabetes mellitus with diabetic cataract

Read Code	Description
C10FE11	Type II diabetes mellitus with diabetic cataract
C10FF00	Type 2 diabetes mellitus with peripheral angiopathy
C10FF11	Type II diabetes mellitus with peripheral angiopathy
C10FG00	Type 2 diabetes mellitus with arthropathy
C10FG11	Type II diabetes mellitus with arthropathy
C10FH00	Type 2 diabetes mellitus with neuropathic arthropathy
C10FH11	Type II diabetes mellitus with neuropathic arthropathy
C10FJ00	Insulin treated Type 2 diabetes mellitus
C10FJ11	Insulin treated Type II diabetes mellitus
C10FK00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus
C10FK11	Hyperosmolar non-ketotic state in type II diabetes mellitus
C10FL00	Type 2 diabetes mellitus with persistent proteinuria
C10FL11	Type II diabetes mellitus with persistent proteinuria
C10FM00	Type 2 diabetes mellitus with persistent microalbuminuria
C10FM11	Type II diabetes mellitus with persistent microalbuminuria
C10FN00	Type 2 diabetes mellitus with ketoacidosis
C10FN11	Type II diabetes mellitus with ketoacidosis
C10FP00	Type 2 diabetes mellitus with ketoacidotic coma
C10FP11	Type II diabetes mellitus with ketoacidotic coma
C10FQ00	Type 2 diabetes mellitus with exudative maculopathy
C10FQ11	Type II diabetes mellitus with exudative maculopathy
C10FR00	Type 2 diabetes mellitus with gastroparesis
C10FR11	Type II diabetes mellitus with gastroparesis
C10FS00	Maternally inherited diabetes mellitus
C10G.00	Secondary pancreatic diabetes mellitus
C10G000	Secondary pancreatic diabetes mellitus without complication
C10H.00	Diabetes mellitus induced by non-steroid drugs
C10H000	Diabetes Mellitus induced by non-steroid drugs without complication
C10M.00	Lipoatrophic diabetes mellitus
C10M000	Lipoatrophic diabetes mellitus without complication
C10N.00	Secondary diabetes mellitus
C10N000	Secondary diabetes mellitus without complication
C10N100	Cystic fibrosis related diabetes mellitus
C10P.00	Diabetes mellitus in remission
C10P100	Type II diabetes mellitus in remission
C10P111	Type 2 diabetes mellitus in remission
PKyP.00	Diabetes insipidus,diabetes mellitus,optic atrophy and deafness

Read code used to identify cases of CKD

Read code	Description
1Z12.00	Chronic kidney disease stage 3
1Z13.00	Chronic kidney disease stage 4
1Z14.00	Chronic kidney disease stage 5
1Z15.00	Chronic kidney disease stage 3A
1Z16.00	Chronic kidney disease stage 3B
1Z1B.00	Chronic kidney disease stage 3 with proteinuria
1Z1B.11	CKD stage 3 with proteinuria
1Z1C.00	Chronic kidney disease stage 3 without proteinuria
1Z1C.11	CKD stage 3 without proteinuria
1Z1D.00	Chronic kidney disease stage 3A with proteinuria
1Z1D.11	CKD stage 3A with proteinuria
1Z1E.00	Chronic kidney disease stage 3A without proteinuria
1Z1E.11	CKD stage 3A without proteinuria
1Z1F.00	Chronic kidney disease stage 3B with proteinuria
1Z1F.11	CKD stage 3B with proteinuria
1Z1G.00	Chronic kidney disease stage 3B without proteinuria
1Z1G.11	CKD stage 3B without proteinuria

Read code	Description
1Z1H.00	Chronic kidney disease stage 4 with proteinuria
1Z1H.11	CKD stage 4 with proteinuria
1Z1J.00	Chronic kidney disease stage 4 without proteinuria
1Z1J.11	CKD stage 4 without proteinuria
1Z1K.00	Chronic kidney disease stage 5 with proteinuria
1Z1K.11	CKD stage 5 with proteinuria
1Z1L.00	Chronic kidney disease stage 5 without proteinuria
1Z1L.11	CKD stage 5 without proteinuria
1Z1N.	CKD with GFR category G1 & albuminuria category A2
1Z1P.	CKD with GFR category G1 & albuminuria category A3
1Z1R.	CKD with GFR category G2 & albuminuria category A2
1Z1S.	CKD with GFR category G2 & albuminuria category A3
1Z1T.	CKD with GFR category G3a & albuminuria category A1
1Z1V.	CKD with GFR category G3a & albuminuria category A2
1Z1W.	CKD with GFR category G3a & albuminuria category A3
1Z1X.	CKD with GFR category G3b & albuminuria category A1
1Z1Y.	CKD with GFR category G3b & albuminuria category A2
1Z1Z.	CKD with GFR category G3b & albuminuria category A3
1Z1a.	CKD with GFR category G4 & albuminuria category A1
1Z1b.	CKD with GFR category G4 & albuminuria category A2
1Z1c.	CKD with GFR category G4 & albuminuria category A3
1Z1d.	CKD with GFR category G5 & albuminuria category A1
1Z1e.	CKD with GFR category G5 & albuminuria category A2
1Z1f.	CKD with GFR category G5 & albuminuria category A3
K053	CKD (Stage 3)
K054	CKD (Stage 4)
K055	CKD (Stage 5)
1Z10.00	Chronic kidney disease stage 1
1Z11.00	Chronic kidney disease stage 2
1Z17.11	CKD stage 1 with proteinuria
1Z18.00	Chronic kidney disease stage 1 without proteinuria
1Z18.11	CKD stage 1 without proteinuria
1Z19.00	Chronic kidney disease stage 2 with proteinuria
1Z19.11	CKD stage 2 with proteinuria
1Z1A.00	Chronic kidney disease stage 2 without proteinuria
1Z1A.11	CKD stage 2 without proteinuria
K051.00	Chronic kidney disease stage 1
K052.00	Chronic kidney disease stage 2
1Z1M.	CKD with GFR category G1 & albuminuria category A1
46TC.00	Urine albumin: creatinine ratio
44lD.00	Urine protein: creatinine ratio
44J7.00	Albumin: creatinine ratio

Read codes used to identify cases of atrial fibrillation

Read code	Description
G573.00	Atrial fibrillation and flutter
G573000	Atrial fibrillation
G573200	Paroxysmal atrial fibrillation
G573300	Non-rheumatic atrial fibrillation
G573400	Permanent atrial fibrillation
G573500	Persistent atrial fibrillation
G573z00	Atrial fibrillation and flutter NOS

Abbreviation: NOS, not otherwise specified; GFR, glomerular filtration rate

Read codes used to identify cases with hypertension

Read code	Description
G2...00	Hypertensive disease
G20..00	Essential hypertension
G20..11	High blood pressure
G20..12	Primary hypertension
G200.00	Malignant essential hypertension
G201.00	Benign essential hypertension
G202.00	Systolic hypertension
G203.00	Diastolic hypertension
G20z.00	Essential hypertension NOS
G20z.11	Hypertension NOS
G21..00	Hypertensive heart disease
G210.00	Malignant hypertensive heart disease
G210000	Malignant hypertensive heart disease without CCF
G210100	Malignant hypertensive heart disease with CCF
G210z00	Malignant hypertensive heart disease NOS
G211.00	Benign hypertensive heart disease
G211000	Benign hypertensive heart disease without CCF
G211100	Benign hypertensive heart disease with CCF
G211z00	Benign hypertensive heart disease NOS
G21z.00	Hypertensive heart disease NOS
G21z000	Hypertensive heart disease NOS without CCF
G21z011	Cardiomegaly - hypertensive
G21z100	Hypertensive heart disease NOS with CCF
G21zz00	Hypertensive heart disease NOS
G22..00	Hypertensive renal disease
G22..11	Nephrosclerosis
G220.00	Malignant hypertensive renal disease
G221.00	Benign hypertensive renal disease
G222.00	Hypertensive renal disease with renal failure
G22z.00	Hypertensive renal disease NOS
G2 2z.11	Renal hypertension
G23..00	Hypertensive heart and renal disease
G230.00	Malignant hypertensive heart and renal disease
G231.00	Benign hypertensive heart and renal disease
G232.00	Hypertensive heart&renal dis wth (congestive) heart failure
G233.00	Hypertensive heart and renal disease with renal failure
G234.00	Hyperten heart&renal dis+both(congestv)heart and renal fail
G23z.00	Hypertensive heart and renal disease NOS
G24..00	Secondary hypertension
G240.00	Secondary malignant hypertension
G240z00	Secondary malignant hypertension NOS
G241.00	Secondary benign hypertension
G241z00	Secondary benign hypertension NOS
G244.00	Hypertension secondary to endocrine disorders
G24z.00	Secondary hypertension NOS
G24z000	Secondary renovascular hypertension NOS
G24zz00	Secondary hypertension NOS
G25..00	Stage 1 hypertension (NICE - Nat Ins for Hth Clin Excl 2011)
G25..11	Stage 1 hypertension
G250.00	Stage 1 hyperten (NICE 2011) without evidence end organ damage
G251.00	Stage 1 hyperten (NICE 2011) with evidence end organ damage
G26..00	Severe hypertension (Nat Inst for Health Clinical Excl 2011)
G26..11	Severe hypertension
G28..00	Stage 2 hypertension (NICE - Nat Ins for Hth Clin Excl 2011)
G2y..00	Other specified hypertensive disease
G2z..00	Hypertensive disease NOS
Gyu2.	[X]Hypertensive diseases
Gyu20.	[X]Other secondary hypertension

Abbreviations: NOS, not otherwise specified; CCF, congestive cardiac failure

Read codes used to identify cases of cardiovascular disease

Note: This list includes CHD, PAD, stroke and TIA combined to create a single CVD group.

Read code	Description
CHD codes	
G3...00	Ischaemic heart disease
G3...11	Arteriosclerotic heart disease
G3...12	Atherosclerotic heart disease
G3...13	IHD - Ischaemic heart disease
G30..00	Acute myocardial infarction
G30..11	Attack - heart
G30..12	Coronary thrombosis
G30..13	Cardiac rupture following myocardial infarction (MI)
G30..14	Heart attack
G30..15	MI - acute myocardial infarction
G30..16	Thrombosis - coronary
G30..17	Silent myocardial infarction
G300.00	Acute anterolateral infarction
G301.00	Other specified anterior myocardial infarction
G301000	Acute anteroapical infarction
G301100	Acute anteroseptal infarction
G301z00	Anterior myocardial infarction NOS
G302.00	Acute inferolateral infarction
G303.00	Acute inferoposterior infarction
G304.00	Posterior myocardial infarction NOS
G305.00	Lateral myocardial infarction NOS
G306.00	True posterior myocardial infarction
G307.00	Acute subendocardial infarction
G307000	Acute non-Q wave infarction
G307100	Acute non-ST segment elevation myocardial infarction
G308.00	Inferior myocardial infarction NOS
G309.00	Acute Q-wave infarct
G30B.00	Acute posterolateral myocardial infarction
G30X.00	Acute transmural myocardial infarction of unspecif site
G30X000	Acute ST segment elevation myocardial infarction
G30y.00	Other acute myocardial infarction
G30y000	Acute atrial infarction
G30y100	Acute papillary muscle infarction
G30y200	Acute septal infarction
G30yz00	Other acute myocardial infarction NOS
G30z.00	Acute myocardial infarction NOS
G31..00	Other acute and subacute ischaemic heart disease
G310.11	Dressler's syndrome
G311.00	Preinfarction syndrome
G311.11	Crescendo angina
G311.12	Impending infarction

Read code	Description
G311.13	Unstable angina
G311.14	Angina at rest
G311000	Myocardial infarction aborted
G311011	MI - myocardial infarction aborted
G311100	Unstable angina
G311200	Angina at rest
G311300	Refractory angina
G311400	Worsening angina
G311500	Acute coronary syndrome
G311z00	Preinfarction syndrome NOS
G312.00	Coronary thrombosis not resulting in myocardial infarction
G31y.00	Other acute and subacute ischaemic heart disease
G31y000	Acute coronary insufficiency
G31y100	Microinfarction of heart
G31y200	Subendocardial ischaemia
G31y300	Transient myocardial ischaemia
G31yz00	Other acute and subacute ischaemic heart disease NOS
G32..00	Old myocardial infarction
G32..11	Healed myocardial infarction
G32..12	Personal history of myocardial infarction
G33..00	Angina pectoris
G330.00	Angina decubitus
G330000	Nocturnal angina
G330z00	Angina decubitus NOS
medcode	description
G33z.00	Angina pectoris NOS
G33z000	Status anginosus
G33z100	Stenocardia
G33z200	Syncope anginosa
G33z300	Angina on effort
G33z400	Ischaemic chest pain
G33z500	Post infarct angina
G33z600	New onset angina
G33z700	Stable angina
G33zz00	Angina pectoris NOS
G34..00	Other chronic ischaemic heart disease
G340.00	Coronary atherosclerosis
G340.11	Triple vessel disease of the heart
G340.12	Coronary artery disease
G340000	Single coronary vessel disease
G340100	Double coronary vessel disease
medcode	description
G342.00	Atherosclerotic cardiovascular disease
G343.00	Ischaemic cardiomyopathy
G344.00	Silent myocardial ischaemia

Read code	Description
G34y.00	Other specified chronic ischaemic heart disease
G34y000	Chronic coronary insufficiency
G34y100	Chronic myocardial ischaemia
G34yz00	Other specified chronic ischaemic heart disease NOS
G34z.00	Other chronic ischaemic heart disease NOS
G34z000	Asymptomatic coronary heart disease
G35..00	Subsequent myocardial infarction
G350.00	Subsequent myocardial infarction of anterior wall
G351.00	Subsequent myocardial infarction of inferior wall
G353.00	Subsequent myocardial infarction of other sites
G35X.00	Subsequent myocardial infarction of unspecified site
G38..00	Postoperative myocardial infarction
G380.00	Postoperative transmural myocardial infarction anterior wall
G381.00	Postoperative transmural myocardial infarction inferior wall
G382.00	Postoperative transmural myocardial infarction other sites
G383.00	Postoperative transmural myocardial infarction unspec site
G384.00	Postoperative subendocardial myocardial infarction
G38z.00	Postoperative myocardial infarction, unspecified
G39..00	Coronary microvascular disease
G3y..00	Other specified ischaemic heart disease
G3z..00	Ischaemic heart disease NOS
Gyu3.00	[X]Ischaemic heart diseases
Gyu3000	[X]Other forms of angina pectoris
Gyu3200	[X]Other forms of acute ischaemic heart disease
Gyu3300	[X]Other forms of chronic ischaemic heart disease
Gyu3400	[X]Acute transmural myocardial infarction of unspecif site
Gyu3500	[X]Subsequent myocardial infarction of other sites
Gyu3600	[X]Subsequent myocardial infarction of unspecified site
PAD codes	
G73..00	Other peripheral vascular disease
G73..11	Peripheral ischaemic vascular disease
G73..12	Ischaemia of legs
G73..13	Peripheral ischaemia
G730.00	Raynaud's syndrome
G730000	Raynaud's disease
G730100	Raynaud's phenomenon
G730111	Vibratory white finger
G730z00	Raynaud's syndrome NOS
G731.00	Thromboangiitis obliterans
G731000	Buerger's disease
G731100	Presenile gangrene
G731z00	Thromboangiitis obliterans NOS
G732.00	Peripheral gangrene
G732000	Gangrene of toe
G732100	Gangrene of foot

Read code	Description
G732200	Gangrene of finger
G732300	Gangrene of thumb
G732400	Gangrene of hand
G733.00	Ischaemic foot
G734.00	Peripheral arterial disease
G735.00	HAVS - Hand-arm vibration syndrome
G735.11	Vibration white finger
G73y.00	Other specified peripheral vascular disease
G73y000	Diabetic peripheral angiopathy
G73y100	Peripheral angiopathic disease EC NOS
G73y200	Acrocyanosis
G73y400	Acroparaesthesia - Schultze's type
G73y411	Schultze's simple acroparaesthesia
G73y500	Acroparaesthesia - Nothnagel's type
G73y511	Nothnagel's vasomotor acroparaesthesia
G73y600	Acroparaesthesia - unspecified
G73y700	Erythrocyanosis
G73y800	Erythromelalgia
G73y811	Erythralgia
G73yz00	Other specified peripheral vascular disease NOS
G73z.00	Peripheral vascular disease NOS
G73z000	Intermittent claudication
G73z011	Claudication
G73z012	Vascular claudication
G73zz00	Peripheral vascular disease NOS
Gyu7400	[X]Other specified peripheral vascular diseases
G734.00	Peripheral arterial disease
G73y.00	Other specified peripheral vascular disease
Stroke/TIA codes	
G61..00	Intracerebral haemorrhage
G61..11	CVA - cerebrovascular accid due to intracerebral haemorrhage
G61..12	Stroke due to intracerebral haemorrhage
G610.00	Cortical haemorrhage
G611.00	Internal capsule haemorrhage
G612.00	Basal nucleus haemorrhage
G613.00	Cerebellar haemorrhage
G614.00	Pontine haemorrhage
G615.00	Bulbar haemorrhage
G616.00	External capsule haemorrhage
G618.00	Intracerebral haemorrhage, multiple localized
G619.00	Lobar cerebral haemorrhage
G61X.00	Intracerebral haemorrhage in hemisphere, unspecified
G61X000	Left sided intracerebral haemorrhage, unspecified
G61X100	Right sided intracerebral haemorrhage, unspecified
G61z.00	Intracerebral haemorrhage NOS
G63y.00	Other precerebral artery occlusion

Read code	Description
G63y000	Cerebral infarct due to thrombosis of precerebral arteries
G63y100	Cerebral infarction due to embolism of precerebral arteries
G64..00	Cerebral arterial occlusion
G64..11	CVA - cerebral artery occlusion
G64..12	Infarction - cerebral
G64..13	Stroke due to cerebral arterial occlusion
G640.00	Cerebral thrombosis
G640000	Cerebral infarction due to thrombosis of cerebral arteries
G641.00	Cerebral embolism
G641.11	Cerebral embolus
G641000	Cerebral infarction due to embolism of cerebral arteries
G64z.00	Cerebral infarction NOS
G64z.11	Brainstem infarction NOS
G64z.12	Cerebellar infarction
G64z000	Brainstem infarction
G64z100	Wallenberg syndrome
G64z111	Lateral medullary syndrome
G64z200	Left sided cerebral infarction
G64z300	Right sided cerebral infarction
G64z400	Infarction of basal ganglia
G66..00	Stroke and cerebrovascular accident unspecified
G66..11	CVA unspecified
G66..12	Stroke unspecified
G66..13	CVA - Cerebrovascular accident unspecified
G660.00	Middle cerebral artery syndrome
G661.00	Anterior cerebral artery syndrome
G662.00	Posterior cerebral artery syndrome
G663.00	Brain stem stroke syndrome
G664.00	Cerebellar stroke syndrome
G665.00	Pure motor lacunar syndrome
G666.00	Pure sensory lacunar syndrome
G667.00	Left sided CVA
G668.00	Right sided CVA
G676000	Cereb infarct due cerebral venous thrombosis, nonpyogenic
G6W..00	Cereb infarct due unsp occlus/stenos precerebr arteries
G6X..00	Cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
Gy62.00	Rupture of dialysis arteriovenous shunt
Gyu..00	[X]Additional circulatory system disease classificatn terms
Gyu0.00	[X]Acute rheumatic fever
Gyu0000	[X]Other acute rheumatic heart disease
Gyu1.00	[X] Chronic rheumatic heart disease
Gyu1000	[X]Other mitral valve diseases
Gyu1100	[X]Other rheumatic aortic valve diseases
Gyu1200	[X]Other tricuspid valve diseases
Gyu1300	[X]Other multiple valve diseases

Read code	Description
Gyu1400	[X]Other specified rheumatic heart diseases
Gyu1500	[X]Multiple valve disease, unspecified
Gyu2.00	[X]Hypertensive diseases
Gyu2000	[X]Other secondary hypertension
Gyu2100	[X]Hypertension secondary to other renal disorders
Gyu3.00	[X]Ischaemic heart diseases
Gyu3000	[X]Other forms of angina pectoris
Gyu3100	[X]Other current complications following acute myocardial infarct
Gyu3200	[X]Other forms of acute ischaemic heart disease
Gyu3300	[X]Other forms of chronic ischaemic heart disease
Gyu3400	[X]Acute transmural myocardial infarction of unspecified site
Gyu3500	[X]Subsequent myocardial infarction of other sites
Gyu3600	[X]Subsequent myocardial infarction of unspecified site
Gyu4.00	[X]Pulmonary heart disease & diseases of pulmonary circulation
Gyu4000	[X]Other specified pulmonary heart diseases
Gyu4100	[X]Other diseases of pulmonary vessels
Gyu5.00	[X]Other forms of heart disease
Gyu5000	[X]Other forms of acute pericarditis
Gyu5100	[X]Other specified diseases of pericardium
Gyu5200	[X]Pericarditis in bacterial diseases classified elsewhere
Gyu5300	[X]Pericarditis in other infectious+parasitic diseases CE
Gyu5400	[X]Pericarditis in other diseases classified elsewhere
Gyu5500	[X]Other nonrheumatic mitral valve disorders
Gyu5600	[X]Other aortic valve disorders
Gyu5700	[X]Other nonrheumatic tricuspid valve disorders
Gyu5800	[X]Other pulmonary valve disorders
Gyu5900	[X]Mitral valve disorders in diseases classified elsewhere
Gyu5A00	[X]Aortic valve disorders in diseases classified elsewhere
Gyu5a00	[X]Other specified cardiac arrhythmias
Gyu5b00	[X]Other ill-defined heart diseases
Gyu5B00	[X]Tricuspid valve disorders/diseases CE
Gyu5c00	[X]Other heart disorders in bacterial diseases CE
Gyu5C00	[X]Pulmonary valve disorders in diseases CE
Gyu5D00	[X]Multiple valve disorders/diseases CE
Gyu5d00	[X]Oth heart disorders/oth infectious+parasitic diseases CE
Gyu5e00	[X]Other heart disorders in other diseases CE
Gyu5E00	[X]Endocarditis,valve unspecified,in diseases CE
Gyu5f00	[X]Nonrheumatic tricuspid valve disorder, unspecified
Gyu5F00	[X]Other acute myocarditis
Gyu5G00	[X]Acute myocarditis, unspecified
Gyu5g00	[X]Cardiovascular disease, unspecified
Gyu5H00	[X]Myocarditis in bacterial diseases classified elsewhere
Gyu5J00	[X]Myocarditis in viral diseases classified elsewhere
Gyu5K00	[X]Myocarditis in other infectious+parasitic diseases CE
Gyu5L00	[X]Myocarditis in other diseases classified elsewhere

Read code	Description
Gyu5M00	[X]Other hypertrophic cardiomyopathy
Gyu5N00	[X]Other restrictive cardiomyopathy
Gyu5P00	[X]Other cardiomyopathies
Gyu5Q00	[X]Cardiomyopathy in infectious+parasitic diseases CE
Gyu5R00	[X]Cardiomyopathy in metabolic diseases CE
Gyu5S00	[X]Cardiomyopathy in nutritional diseases CE
Gyu5T00	[X]Cardiomyopathy in other diseases classified elsewhere
Gyu5U00	[X]Other and unspecified atrioventricular block
Gyu5V00	[X]Other and unspecified fascicular block
Gyu5W00	[X]Other and unspecified right bundle-branch block
Gyu5X00	[X]Other specified heart block
Gyu5Y00	[X]Other specified conduction disorders
Gyu5Z00	[X]Other and unspecified premature depolarization
Gyu6.00	[X]Cerebrovascular diseases
Gyu6000	[X]Subarachnoid haemorrhage from other intracranial arteries
Gyu6100	[X]Other subarachnoid haemorrhage
Gyu6200	[X]Other intracerebral haemorrhage
Gyu6300	[X]Cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
Gyu6400	[X]Other cerebral infarction
Gyu6500	[X]Occlusion and stenosis of other precerebral arteries
Gyu6600	[X]Occlusion and stenosis of other cerebral arteries
Gyu6F00	[X]Intracerebral haemorrhage in hemisphere, unspecified
Gyu6G00	[X]Cereb infarct due unsp occlus/stenos precerebr arteries
G65..00	Transient cerebral ischaemia
G65..11	Drop attack
G65..12	Transient ischaemic attack
G65..13	Vertebro-basilar insufficiency
G650.00	Basilar artery syndrome
G650.11	Insufficiency - basilar artery
G651.00	Vertebral artery syndrome
G651000	Vertebro-basilar artery syndrome
G652.00	Subclavian steal syndrome
G653.00	Carotid artery syndrome hemispheric
G654.00	Multiple and bilateral precerebral artery syndromes
G656.00	Vertebrobasilar insufficiency
G657.00	Carotid territory transient ischaemic attack
G65y.00	Other transient cerebral ischaemia
G65z.00	Transient cerebral ischaemia NOS
G65z000	Impending cerebral ischaemia
G65z100	Intermittent cerebral ischaemia
G65zz00	Transient cerebral ischaemia NOS
ZV12D00	[V]Personal history of transient ischaemic attack
Fyu5500	[X]Other transnt cerebral ischaemic attacks+related syndroms

Abbreviations: CHD, coronary heart disease; NOS, not otherwise specified; [X], Cross referenced to specific ICD-10 codes; PAD, peripheral arterial disease; TIA, transient ischaemic attack; CVA, cardiovascular accident

Read codes used to identify cases of heart failure

Read code	Description
G58..00	Heart failure
G58..11	Cardiac failure
G580.00	Congestive heart failure
G580.11	Congestive cardiac failure
G580.12	Right heart failure
G580.13	Right ventricular failure
G580.14	Biventricular failure
G580000	Acute congestive heart failure
G580100	Chronic congestive heart failure
G580200	Decompensated cardiac failure
G580300	Compensated cardiac failure
G580400	Congestive heart failure due to valvular disease
G581.00	Left ventricular failure
G581.11	Asthma - cardiac
G581.12	Pulmonary oedema - acute
G581.13	Impaired left ventricular function
G581000	Acute left ventricular failure
G582.00	Acute heart failure
G583.00	Heart failure with normal ejection fraction
G583.11	HFNEF - heart failure with normal ejection fraction
G583.12	Heart failure with preserved ejection fraction
G584.00	Right ventricular failure
G58z.00	Heart failure NOS
G58z.11	Weak heart
G58z.12	Cardiac failure NOS
G1yz100	Rheumatic left ventricular failure
662f.00	New York Heart Association classification - class I
662F.00	Hypertension treatment. started
662G.00	Hypertensive treatment changed
662g.00	New York Heart Association classification - class II
662H.00	Hypertension treatment stopped
662h.00	New York Heart Association classification - class III
662i.00	New York Heart Association classification - class IV
585f.00	Echocardiogram shows left ventricular systolic dysfunction
G5yy900	Left ventricular systolic dysfunction
G5yyD00	Left ventricular cardiac dysfunction

Abbreviations: NOS, not otherwise specified

Drug codes used to identify prescription of non-steroidal anti-inflammatory drugs

Drug code	Generic name
52437979	Naproxen 250mg tablets
52438979	Mefenamic acid 500mg tablets
53244979	Ibuprofen lysine 400mg oral powder sachets
53245979	Ibuprofen lysine 400mg oral powder sachets
54501979	Ibuprofen 200mg capsules
59408979	Indometacin 25mg capsules
60599979	Etoricoxib 30mg tablets
68279979	Naproxen 75mg/5ml oral suspension
68284979	Naproxen 100mg/5ml oral suspension
70267978	Ibuprofen 200mg tablets
79388978	Misoprostol 400microgram tablets
79449979	Naproxen 500mg/5ml oral suspension
79451979	Naproxen 250mg/5ml oral suspension
79837979	Indometacin 25mg/5ml oral solution
81659998	Ibuprofen 200mg effervescent tablets
81664998	Ibuprofen 10% gel
81960998	Ketoprofen 200mg modified-release capsules
82031998	Celecoxib 400mg capsules
82272998	Ibuprofen 5% spray
82307998	Tenoxicam 20mg tablets
82366998	Ibuprofen lysine 200mg tablets
82686998	Ibuprofen 5% gel
82872978	Meloxicam 7.5mg orodispersible tablets sugar free
82874978	Meloxicam 15mg orodispersible tablets sugar free
82923998	Ketoprofen 150mg modified-release capsules
82924998	Ketoprofen 200mg modified-release capsules
82925998	Ketoprofen 100mg modified-release capsules
82926998	Ketoprofen 100mg suppositories
82927998	Ketoprofen 100mg capsules
82928998	Ketoprofen 50mg capsules
83070998	Naproxen 200mg/5ml oral suspension
83154998	Ibuprofen 400mg capsules
83155998	Ibuprofen 200mg capsules
83433998	Ibuprofen 400mg tablets
83434998	Aspirin 500mg effervescent tablets sugar free
83445998	Piroxicam 20mg orodispersible tablets sugar free
83471998	Ibuprofen 800mg modified-release tablets
83494998	Ibuprofen sodium dihydrate 200mg tablets
83497998	Ibuprofen sodium dihydrate 200mg tablets
83500998	Ibuprofen 200mg tablets
83608998	Etodolac 600mg modified-release tablets
83611998	Flurbiprofen 200mg modified release capsules
83984998	Naproxen 250mg gastro-resistant tablets

Drug code	Generic name
84019998	Ibuprofen 5% gel
84058998	Ibuprofen 5% cream
84104998	Ibuprofen 200mg/5ml oral suspension
84128998	Etoricoxib 30mg tablets
84129998	Etoricoxib 30mg tablets
84153998	Ibuprofen 400mg tablets
84155998	Ibuprofen 200mg tablets
84160998	Ibuprofen 200mg tablets
84174998	Ibuprofen 5% gel
84433998	Ibuprofen 200mg capsules
84434998	Ibuprofen lysine 400mg tablets
84435998	Ibuprofen 400mg capsules
84437998	Ibuprofen lysine 200mg tablets
84490998	Ibuprofen 200mg tablets
84553998	Ibuprofen 200mg capsules
84554998	Ibuprofen 200mg tablets
84555998	Ibuprofen 200mg tablets
84973998	Indometacin oral solution
85154998	Diclofenac sodium 25mg gastro-resistant tablets
85272998	Naproxen oral solution
85428998	Etodolac 600mg modified-release tablets
85741998	Ibuprofen 400mg capsules
85783998	Ibuprofen 400mg capsules
85891979	Ibuprofen lysine 200mg tablets
85953998	Phenylbutazone 100mg tablets
86170998	Ibuprofen lysine 400mg tablets
86171998	Ibuprofen 400mg tablets
86329998	Ibuprofen 10mg/2ml solution for infusion ampoules
86594998	Ibuprofen lysine 200mg tablets
86624998	Dexibuprofen 400mg tablets
86628998	Dexibuprofen 300mg tablets
86629998	Dexibuprofen 400mg tablets
86635998	Dexibuprofen 300mg tablets
86940998	Ibuprofen 5% gel
86953998	Ibuprofen 200mg tablets
87070998	Ibuprofen 300mg modified-release capsules
87413998	Ketoprofen 50mg capsules
87936998	Aspirin 500mg granules sachets sugar free
88047997	Ketoprofen 200mg modified-release capsules
88047998	Ketoprofen 100mg modified-release capsules
88138998	Naproxen 500mg gastro-resistant tablets
88139998	Naproxen 250mg gastro-resistant tablets
88143998	Ibuprofen 5% gel
88145996	Ibuprofen 400mg tablets
88145998	Ibuprofen 5% gel

Drug code	Generic name
88174998	Ibuprofen 200mg capsules
88204998	Ibuprofen 5% mousse
88205998	Ibuprofen 5% foam
88228998	Ibuprofen lysine 200mg tablets
88233998	Ibuprofen 200mg tablets
88284998	Ibuprofen 400mg tablets
88442998	Mefenamic acid 250mg capsules
88455998	Indometacin 75mg modified-release capsules
88527997	Ibuprofen 200mg capsules
88527998	Ibuprofen 200mg tablets
88817998	Tolfenamic acid 200mg tablets
88943998	Ketoprofen 200mg modified release capsules
88970998	Etoricoxib 60mg tablets
88977998	Etoricoxib 120mg tablets
89014998	Ibuprofen 200mg modified-release capsules
89117998	Ibuprofen 200mg tablets
89137997	Ketoprofen 200mg modified release capsules
89139998	Ibuprofen 5% gel
89217998	Aspirin 300mg gastro-resistant tablets
89398979	Etoricoxib 90mg tablets
89404979	Etoricoxib 60mg tablets
89472997	Ibuprofen 5% gel
89479998	Tolfenamic acid 200mg capsules
89484997	Tolfenamic acid 200mg tablets
89484998	Tolfenamic acid 200mg capsule
89499998	Ibuprofen 10% gel
89572998	Etodolac 300mg capsules
89580998	Ibuprofen 200mg tablets
89593997	Ibuprofen 400mg tablets
89621998	Ibuprofen 200mg orodispersible tablets sugar free
89691997	Ibuprofen 400mg tablets
89691998	Ibuprofen 200mg tablets
89760998	Ibuprofen 300mg modified-release capsules
89801998	Flurbiprofen 8.75mg lozenges
89890998	Ibuprofen lysine 200mg tablets
89898998	Codeine 8mg with aspirin 500mg soluble tablets
89909998	Ketoprofen 200mg modified release capsules
89966998	Ketoprofen 2.5% gel
89993998	Ibuprofen 5% gel
90116998	Ibuprofen 200mg tablets
90125998	Ibuprofen 200mg tablets
90278998	Aspirin 500mg effervescent tablets sugar free
90351997	Meloxicam 15mg tablets
90351998	Meloxicam 7.5mg tablets
90361997	Meloxicam 15mg tablets

Drug code	Generic name
90361998	Meloxicam 7.5mg tablets
90368997	Celecoxib 200mg capsules
90368998	Celecoxib 100mg capsules
90377997	Aspirin 500mg effervescent tablets sugar free
90377998	Aspirin 300mg effervescent tablets sugar free
90635998	Aceclofenac 100mg tablets
90636998	Aceclofenac 100mg tablets
90709998	Flurbiprofen 8.75mg lozenges
90846998	Ketoprofen 200mg modified-release capsules
90869998	Ibuprofen 400mg tablets
90954998	Flurbiprofen 8.75mg lozenges
91081996	Ibuprofen 200mg tablets
91081997	Ibuprofen 200mg capsules
91081998	Ibuprofen 400mg granules
91105998	Indometacin 25mg modified-release capsules
91109997	Mefenamic acid 500mg tablets
91109998	Mefenamic acid 250mg capsules
91120998	Ibuprofen 200mg tablets
91155998	Ibuprofen 10% gel
91315998	Ketoprofen 2.5% gel
91421998	Dexketoprofen 25mg tablets
91438979	Ibuprofen 10% gel
91442979	Piroxicam 0.5% gel
91443979	Piroxicam 0.5% gel
91446979	Piroxicam 0.5% gel
91447979	Piroxicam 0.5% gel
91451979	Piroxicam 0.5% gel
91463997	Piroxicam 20mg capsules
91463998	Piroxicam 10mg capsules
91466979	Ketoprofen 2.5% gel
91475979	Ibuprofen 5% gel
91479979	Celecoxib 100mg capsules
91486979	Celecoxib 200mg capsules
91502998	Indometacin 75mg modified release capsules
91517998	Piroxicam 0.5% gel
91523998	Ibuprofen 10% gel
91581997	Celecoxib 200mg capsules
91581998	Celecoxib 100mg capsules
91682979	Tenoxicam 20mg tablets
91713979	Piroxicam 20mg orodispersible tablets sugar free
91751979	Naproxen 250mg tablets
91752979	Naproxen 250mg tablets
91757979	Naproxen 250mg tablets
91763979	Mefenamic acid 500mg tablets
91770979	Mefenamic acid 250mg capsules

Drug code	Generic name
91774979	Ketoprofen 200mg modified-release capsules
91774998	Indometacin 75mg modified release capsules
91777998	Ibuprofen 10% gel
91778979	Ketoprofen 200mg modified-release capsules
91782979	Ketoprofen 100mg modified-release capsules
91815997	Mefenamic acid 500mg tablets
91815998	Mefenamic acid 250mg capsules
91841998	Aspirin 300mg orodispersible tablets sugar free
91843979	Ibuprofen 400mg tablets
91850979	Ibuprofen 400mg tablets
91851979	Ibuprofen 400mg tablets
91854979	Ibuprofen 400mg tablets
91856979	Ibuprofen 200mg tablets
91863979	Ibuprofen 200mg tablets
91864979	Ibuprofen 200mg tablets
91877998	Ketoprofen 200mg modified release capsules
91920997	Mefenamic acid 500mg tablets
91920998	Mefenamic acid 250mg capsules
91965979	Aspirin 300mg tablets
91988998	Etoricoxib 90mg tablets
91989998	Etoricoxib 60mg tablets
91990998	Etoricoxib 120mg tablets
91991998	Etoricoxib 90mg tablets
92092990	Ketoprofen 2.5% gel
92112998	Ketoprofen 200mg modified-release capsules
92113998	Indometacin 75mg modified release capsules
92158998	Tiaprofenic acid 300mg tablets
92169998	Ibuprofen 200mg capsules
92189998	Piroxicam betadex 20mg tablets
92290998	Ibuprofen 200mg orodispersible tablets sugar free
92550990	Mefenamic acid 500mg tablets
92551990	Mefenamic acid 250mg capsules
92671998	Aspirin 500mg modified-release tablets
92706998	Aspirin 300mg gastro-resistant tablets
92738990	Naproxen 500mg gastro-resistant tablets
92778998	Aspirin 300mg orodispersible tablets sugar free
92801998	Piroxicam 10mg capsules
92851998	Tenoxicam 20mg injection plus diluent
92863998	Etodolac 600mg modified-release tablets
92864998	Etodolac 600mg modified-release tablets
92950996	Naproxen 500mg gastro-resistant tablets
92950997	Naproxen 250mg tablets
92950998	Naproxen 500mg tablets
92953997	Ketoprofen 100mg capsules
92954997	Indometacin 75mg modified release capsules

Drug code	Generic name
92954998	Indometacin 25mg capsules
92958998	Indometacin 1mg injection (powder for reconstitution)
92965998	Ibuprofen 400mg tablets
93029998	Naproxen 500mg modified release tablets
93089997	Ketoprofen 200mg modified-release capsules
93089998	Ketoprofen 100mg modified-release capsules
93099997	Aspirin 300mg modified release tablets
93135996	Naproxen 500mg modified-release tablets
93135997	Naproxen 375mg modified-release tablet
93135998	Naproxen 500mg granules
93152998	Ketoprofen 100mg/2ml injection
93169996	Naproxen 375mg gastro-resistant tablets
93169997	Naproxen 500mg gastro-resistant tablets
93169998	Naproxen 250mg gastro-resistant tablets
93170997	Naproxen 500mg gastro-resistant tablets
93170998	Naproxen 250mg gastro-resistant tablets
93218998	Ibuprofen 200mg tablets
93235990	Ibuprofen 10% gel
93247998	Acemetacin 60mg capsules
93261998	Nabumetone 500mg dispersible tablets sugar free
93267998	Naproxen 375mg tablets
93272996	Ibuprofen lysine 400mg tablets
93272997	Ibuprofen 200mg tablets
93272998	Ibuprofen 200mg tablets
93351990	Indometacin 25mg capsules
93368998	Aspirin 300mg dispersible tablets
93579990	Meloxicam 15mg tablets
93580990	Meloxicam 7.5mg tablets
93625998	Ibuprofen 5% cream
93626996	Ibuprofen 5% spray
93626997	Ibuprofen 5% gel
93626998	Ibuprofen 5% cream
93688992	Aspirin s/r 500 mg tab
93698990	Meloxicam 7.5mg tablets
93726990	Meloxicam 7.5mg tablets
93731992	Aspirin disp 500 mg tab
93756990	Ibuprofen 400mg tablets
93866998	Ibuprofen 200mg tablets
94020992	Aspirin soluble 600 mg tab
94073992	Aspirin 600 mg sup
94152998	Ibuprofen 10% gel
94165992	Fenoprofen 300mg tablets
94213998	Aspirin 600mg tablets
94214997	Aspirin 300mg gastro-resistant tablets
94214998	Aspirin 324mg gastro-resistant tablets

Drug code	Generic name
94215996	Aspirin 300mg orodispersible tablets sugar free
94215997	Aspirin 300mg orodispersible tablets sugar free
94216997	Aspirin 300mg modified-release tablets
94216998	Aspirin 324mg modified-release tablets
94240992	Ibuprofen 200mg tablets
94254998	Aspirin 300mg gastro-resistant tablets
94257998	Piroxicam 0.5% gel
94258997	Piroxicam 0.5% gel
94258998	Piroxicam 0.5% gel
94262998	Aspirin 500mg with papaveretum 7.71mg dispersible tablets
94352998	Ketoprofen 100mg/2ml solution for injection ampoules
94437990	Piroxicam 0.5% gel
94459990	Ibuprofen 400mg tablets
94489996	Ketoprofen 150mg modified-release capsules
94489997	Ketoprofen 200mg modified-release capsules
94489998	Ketoprofen 100mg modified-release capsules
94513997	Aspirin 300mg effervescent tablets
94514996	Tenoxicam 20mg effervescent tablets
94514997	Tenoxicam 20mg/sachet granules
94514998	Tenoxicam 20mg tablets
94515996	Tenoxicam 20mg effervescent tablets
94515998	Tenoxicam 20mg tablets
94589997	Aspirin 300mg effervescent tablets
94589998	Aspirin 100mg effervescent tablets
94607996	Ketorolac 10mg/1ml solution for injection ampoules
94607997	Ketorolac 10mg tablets
94607998	Ketorolac 30mg/1ml solution for injection ampoules
94608996	Ketorolac trometamol 10mg/1ml injection
94608997	Ketorolac trometamol 10mg tablets
94608998	Ketorolac 30mg/1ml solution for injection ampoules
94626990	Ibuprofen 400mg tablets
94631998	Misoprostol 200microgram tablets
94632998	Misoprostol 200microgram vaginal tablets
94651992	Ibuprofen 400mg tablets
94667992	Aspirin m/f 324 mg tab
94668992	Aspirin 325 mg cap
94671992	Aspirin 500 mg sup
94674992	Aspirin disp 600 mg tab
94678990	Ibuprofen 400mg tablets
94678992	Aspirin soluble 400 mg tab
94679992	Aspirin soluble 500 mg tab
94709997	Aspirin 300mg effervescent tablets sugar free
94709998	Aspirin 500mg effervescent tablets sugar free
94743992	Brufen sup
94759998	Aspirin 500mg with cyclizine 25mg effervescent tablets

Drug code	Generic name
94784998	Flurbiprofen 200mg modified-release capsules
94798998	Flurbiprofen 200mg modified release capsules
94805997	Ibuprofen 600mg effervescent granules sachets
94805998	Ibuprofen 800mg modified-release tablets
94809997	Piroxicam 20mg dispersible tablets
94809998	Piroxicam 10mg dispersible tablets
94832996	Indometacin 25mg modified release tablets
94832997	Indometacin 50mg modified release tablets
94832998	Indometacin 75mg modified release tablets
94874998	Ibuprofen 800mg tablets
94875998	Ibuprofen 800mg tablets
94887998	Ibuprofen 200mg orodispersible tablets sugar free
94907997	Nabumetone 500mg/5ml oral suspension sugar free
94907998	Nabumetone 500mg tablets
94914997	Nabumetone 500mg/5ml suspension
94914998	Nabumetone 500mg tablets
94916998	Ibuprofen 5% foam
94928998	Tiaprofenic acid 300mg modified release capsules
95006998	Ketoprofen 2.5% gel
95013998	Ketoprofen 2.5% gel
95014992	Fenoprofen 300mg tablets
95061998	Dexketoprofen 25mg tablets
95075998	Indometacin 25mg modified-release tablets
95093996	Naproxen 500mg gastro-resistant tablets
95093997	Naproxen 375mg gastro-resistant tablets
95093998	Naproxen 250mg gastro-resistant tablets
95143992	Ibuprofen 200mg capsules
95167996	Tiaprofenic acid 300mg sachets
95167997	Tiaprofenic acid 300mg tablets
95167998	Tiaprofenic acid 200mg tablets
95172990	Ibuprofen 10% gel
95191990	Piroxicam 0.5% gel
95212992	Aspirin 500mg modified release tablets
95227997	Sulindac 200mg tablets
95227998	Sulindac 100mg tablets
95340990	Ibuprofen 200mg tablets
95347990	Ibuprofen 400mg tablets
95348990	Ibuprofen 200mg tablets
95351990	Aspirin 300mg dispersible tablets
95496990	Piroxicam 0.5% gel
95496998	Piroxicam 10mg capsules
95497998	Piroxicam 20mg suppositories
95498996	Piroxicam 20mg orodispersible tablets sugar free
95498997	Piroxicam 20mg capsules
95498998	Piroxicam 10mg capsules

Drug code	Generic name
95539997	Phenylbutazone 200mg tablets
95539998	Phenylbutazone 100mg tablets
95540998	Phenylbutazone 100mg gastro-resistant tablets
95541997	Phenylbutazone 200mg tablets
95541998	Phenylbutazone 100mg tablets
95611990	Ketoprofen 2.5% gel
95753998	Naproxen sodium 275mg tablets
95754997	Naproxen 125mg/5ml oral suspension
95754998	Naproxen 500mg suppositories
95909996	Mefenamic acid 50mg/5ml oral suspension
95909997	Mefenamic acid 500mg tablets
95909998	Mefenamic acid 250mg dispersible tablet
95911992	Aspirin 300mg dispersible tablets
95992990	Ibuprofen 5% gel
96035996	Ketoprofen 100mg suppositories
96035997	Ketoprofen 100mg capsules
96035998	Ketoprofen 50mg capsules
96310989	Ibuprofen 400mg tablets
96310990	Ibuprofen 200mg tablets
96369990	Ibuprofen 200mg tablets
96405996	Ibuprofen 600mg tablets
96405997	Ibuprofen 400mg tablets
96405998	Ibuprofen 200mg tablets
96407989	Indometacin 50mg capsules
96407990	Indometacin 25mg capsules
96414990	Aspirin 300mg dispersible tablets
96418989	Ibuprofen 400mg tablets
96451997	Naproxen 500mg tablets
96451998	Naproxen 250mg tablets
96452997	Naproxen 500mg tablets
96452998	Naproxen 250mg tablets
96495996	Flurbiprofen 100mg suppositories
96495997	Flurbiprofen 100mg tablets
96495998	Flurbiprofen 50mg tablets
96555996	Fenoprofen 600mg tablets
96555997	Fenoprofen 300mg tablets
96555998	Fenoprofen 200mg tablet
96566989	Aspirin 300mg dispersible tablets
96566990	Aspirin 300mg tablets
96569992	Aspirin sr 300 mg tab
96583990	Ibuprofen 5% gel
96625988	Naproxen 500mg gastro-resistant tablets
96841990	Piroxicam 20mg dispersible tablets
96842989	Piroxicam 20mg capsules
96851990	Mefenamic acid 250mg capsules

Drug code	Generic name
96879990	Naproxen 250mg gastro-resistant tablets
96921998	Nabumetone 500mg dispersible tablets
96939989	Ibuprofen 400mg tablets
96939990	Ibuprofen 200mg tablets
96961989	Mefenamic acid 500mg tablets
96961990	Mefenamic acid 250mg capsules
97056996	Ibuprofen 600mg tablets
97056997	Ibuprofen 400mg tablets
97056998	Ibuprofen 200mg tablets
97088992	Aspirin 324mg gastro-resistant tablets
97100989	Naproxen 500mg gastro-resistant tablets
97100990	Naproxen 250mg gastro-resistant tablets
97104988	Ibuprofen 600mg tablets
97104989	Ibuprofen 400mg tablets
97104990	Ibuprofen 200mg tablets
97106997	Ibuprofen 400mg tablets
97107997	Ibuprofen 400mg tablets
97111990	Mefenamic acid 250mg capsules
97114990	Indometacin 75mg modified-release capsules
97180989	Naproxen 500mg gastro-resistant tablets
97180990	Naproxen 250mg gastro-resistant tablets
97181990	Aspirin 300mg gastro-resistant tablets
97305992	Dihydrocodeine tartrate/aspirin 300 mg tab
97356998	Indometacin 75mg modified-release capsules
97357996	Indometacin 100mg suppositories
97358997	Ibuprofen 400mg tablets
97358998	Ibuprofen 200mg tablets
97537989	Aspirin 300mg gastro-resistant tablets
97550988	Naproxen 250mg gastro-resistant tablets
97551989	Ibuprofen 400mg tablets
97551990	Ibuprofen 200mg tablets
97564998	Naproxen sodium 275mg tablets
97565998	Naproxen 500mg tablets
97566996	Naproxen 500mg suppositories
97566997	Naproxen 125mg/5ml suspension
97566998	Naproxen 250mg tablets
97593996	Ibuprofen 600mg tablets
97594997	Ibuprofen 400mg tablets
97594998	Ibuprofen 200mg tablets
97641992	Ketorolac 30mg/1ml solution for injection ampoules
97657997	Naproxen 500mg tablets
97657998	Naproxen 250mg tablets
97658997	Naproxen 500mg modified release tablets
97658998	Naproxen 375mg modified release tablets
97668998	Piroxicam 20mg/1ml injection

Drug code	Generic name
97674989	Ibuprofen 400mg tablets
97674990	Ibuprofen 200mg tablets
97678997	Ibuprofen 400mg tablets
97700990	Naproxen 500mg gastro-resistant tablets
97712997	Naproxen 500mg tablets
97712998	Naproxen 250mg tablets
97746990	Ketoprofen 200mg modified-release capsules
97748998	Piroxicam 20mg/1ml solution for injection ampoules
97902990	Piroxicam 20mg dispersible tablets
97906996	Ibuprofen 600mg tablets
97906997	Ibuprofen 400mg tablets
97906998	Ibuprofen 200mg tablets
97918989	Aspirin 300mg dispersible tablets
98040988	Ketoprofen 100mg capsules
98041990	Indometacin 50mg capsules
98127989	Mefenamic acid 500mg tablets
98127990	Mefenamic acid 250mg capsules
98134989	Flurbiprofen 50mg tablets
98137998	Ibuprofen 10% gel
98142989	Aspirin 300mg dispersible tablets
98150990	Ibuprofen 600mg tablets
98151990	Ibuprofen 600mg tablets
98166996	Piroxicam 20mg orodispersible tablets sugar free
98166997	Piroxicam 20mg dispersible tablets
98166998	Piroxicam 10mg dispersible tablets
98280998	Aspirin & metoclopramide 450mg+5mg effervescent tablets
98399998	Indometacin 75mg modified-release capsules
98419997	Aspirin 300mg dispersible tablets
98419998	Aspirin 300mg tablets
98426989	Piroxicam 20mg capsules
98426990	Piroxicam 10mg capsules
98429998	Ibuprofen & codeine phosphate 300mg+20mg modified release tab
98495989	Mefenamic acid 500mg tablets
98495990	Mefenamic acid 250mg capsules
98513990	Aspirin 300mg tablets
98515998	Ibuprofen 600mg tablets
98516998	Ibuprofen 400mg tablets
98528990	Ibuprofen 600mg tablets
98529988	Ibuprofen 600mg tablets
98529989	Ibuprofen 400mg tablets
98529990	Ibuprofen 200mg tablets
98530988	Ibuprofen 600mg tablets
98530989	Ibuprofen 400mg tablets
98530990	Ibuprofen 200mg tablets
98555989	Ibuprofen 400mg tablets

Drug code	Generic name
98555990	Ibuprofen 200mg tablets
98578998	Ibuprofen 300mg modified-release capsules
98592988	Aspirin 300mg dispersible tablets
98600989	Ketoprofen 200mg modified-release capsules
98621989	Ketoprofen 200mg modified-release capsules
98654998	Mefenamic acid 50mg/5ml paediatric suspension
98671988	Indometacin 50mg capsules
98671989	Indometacin 25mg capsules
98671990	Indometacin 75mg modified-release capsules
98672989	Indometacin 25mg capsules
98672990	Indometacin 50mg capsules
98673988	Ibuprofen 200mg tablets
98673989	Ibuprofen 400mg tablets
98673990	Ibuprofen 600mg tablets
98674988	Naproxen 250mg gastro-resistant tablets
98674989	Naproxen 500mg gastro-resistant tablets
98674990	Naproxen 250mg tablets
98693998	Acemetacin 60mg capsules
98758998	Ketoprofen 200mg modified release capsules
98764998	Ibuprofen 200mg modified-release capsules
98779998	Ketoprofen 100mg suppositories
98907998	Naproxen 500mg/sachet granules
99334997	Aspirin 600mg gastro-resistant tablets
99334998	Aspirin 300mg gastro-resistant tablets
99442989	Piroxicam 20mg capsules
99442990	Piroxicam 10mg capsules
99444989	Piroxicam 20mg capsules
99444990	Piroxicam 10mg capsules
99445989	Piroxicam 20mg capsules
99445990	Piroxicam 10mg capsules
99466996	Ibuprofen 400mg tablets
99466997	Ibuprofen 200mg tablets
99482998	Piroxicam betadex 20mg tablets
99516998	Tiaprofenic acid 300mg modified-release capsules
99517989	Mefenamic acid 500mg tablets
99517990	Mefenamic acid 250mg capsules
99519989	Mefenamic acid 500mg tablets
99519990	Mefenamic acid 250mg capsules
99520989	Mefenamic acid 500mg tablets
99520990	Mefenamic acid 250mg capsules
99535996	Indometacin 25mg/5ml sugar free suspension
99535997	Indometacin 50mg capsules
99535998	Indometacin 25mg capsules
99539996	Indometacin 100mg suppositories
99539997	Indometacin 50mg capsules

Drug code	Generic name
99539998	Indometacin 25mg capsules
99550989	Indometacin 50mg capsules
99550990	Indometacin 25mg capsules
99551990	Indometacin 25mg capsules
99552988	Indometacin 100mg suppositories
99552989	Indometacin 50mg capsules
99552990	Indometacin 25mg capsules
99553989	Indometacin 75mg modified-release capsules
99553990	Indometacin 25mg capsules
99557988	Ibuprofen 600mg tablets
99557989	Ibuprofen 400mg tablets
99557990	Ibuprofen 200mg tablets
99558988	Ibuprofen 400mg tablets
99558989	Ibuprofen 600mg tablets
99558990	Ibuprofen 200mg tablets
99621996	Flurbiprofen 100mg suppositories
99621997	Flurbiprofen 100mg tablets
99621998	Flurbiprofen 50mg tablets
99652997	Fenoprofen 600mg tablets
99652998	Fenoprofen 300mg tablets
99653996	Piroxicam 20mg suppositories
99653997	Piroxicam 20mg capsules
99653998	Piroxicam 10mg capsules
99728988	Naproxen 500mg gastro-resistant tablets
99728989	Naproxen 500mg gastro-resistant tablets
99728990	Naproxen 250mg tablets
99730989	Naproxen 500mg gastro-resistant tablets
99730990	Naproxen 250mg tablets
99731989	Naproxen 500mg gastro-resistant tablets
99731990	Naproxen 250mg tablets
99807988	Aspirin 300mg tablets
99807989	Aspirin 300mg dispersible tablets
99808988	Aspirin 300mg tablets
99808989	Aspirin 300mg dispersible tablets
99810989	Aspirin 300mg dispersible tablets
99823997	Sulindac 200mg tablets
99823998	Sulindac 100mg tablets
99824998	Aspirin 300mg tablets
99868998	Ibuprofen 200mg tablets
99876992	Aspirin 325 mg tab

Read code used to detect glycaemic abnormality

Note: Fructosamine is an alternative test of assessing glucose abnormality in patients where HbA1c cannot be reliably measured

Read code	Description
Sickle cell	
D1060	Sickle cell anaemia of unspecific type
D1061	Sickle cell anaemia with no crisis
D1062	Sickle cell anaemia with crisis
D1063	Sickle cell anaemia with haemoglobin C disease
D1064	Sickle cell anaemia with haemoglobin D disease
D1065	Sickle cell anaemia with haemoglobin E disease
D106z	Sickle cell anaemia NOS
Test result code	
42D4	Sickle cell present
Thalassaemia	
D1040	Thalassaemia major NEC
D1041	Thalassaemia minor NEC
D1042	Thalassaemia with haemoglobin S disease
D1043	Alpha Thalassaemia
D1046	Beta intermediate Thalassaemia
D1047	Beta major Thalassaemia
D1048	Beta minor Thalassaemia
D104z	Thalassaemia NOS
Lab test code	
44TD.00	Fructosamine

Abbreviations: NEC; not elsewhere classified

21.0 Appendix 12: THIN Scientific Review Committee approval

SRC Feedback

Researcher Name: Dr Feroz Jadhakhan
Organisation: University of Birmingham
SRC Reference Number: 14-038
Date: 13th June 2014
Study title: Cumulative incidence of Chronic Kidney Disease (CKD) in young adults (aged 18 to 40 years) with Impaired Glucose Tolerance (IGT)
Committee opinion: Approved

The following feedback has been supplied by the SRC.

Notes from the Chair:

Approved

We are pleased to inform that you can proceed with the study as this is now approved. CSD Medical Research will let the relevant Ethics committee know this study has been approved by the SRC.

Once the study has been completed and published, it is important for you to inform CSD Medical Research in order for us to advise the SRC and your reference number to be closed.

References to all published studies are added to our website enabling other researchers to become aware of your work. Copies of publication(s), where available, will be appreciated.

I wish you and your team all the best with the study progression.

Mustafa Dungarwalla
Research Associate

SRC
Scientific Review Committee

22.0 Appendix 13: Comparison of prevalence of prediabetes/non-diabetic hyperglycaemia

Authors	Database	Age	Prevalence of non-diabetic hyperglycaemia
NHS-DPP (2015) (77)	HSE	16 to 39	2.6% in (2009 - 2013)
			Prevalence of pre-diabetes
Mainous et al (2014) (76)	HSE	≥16 years	35.3% in (2011)

Abbreviations: IGR, impaired glucose regulation; IFG, impaired fasting glucose; GPRD, general practice research database; HSE, health survey for England

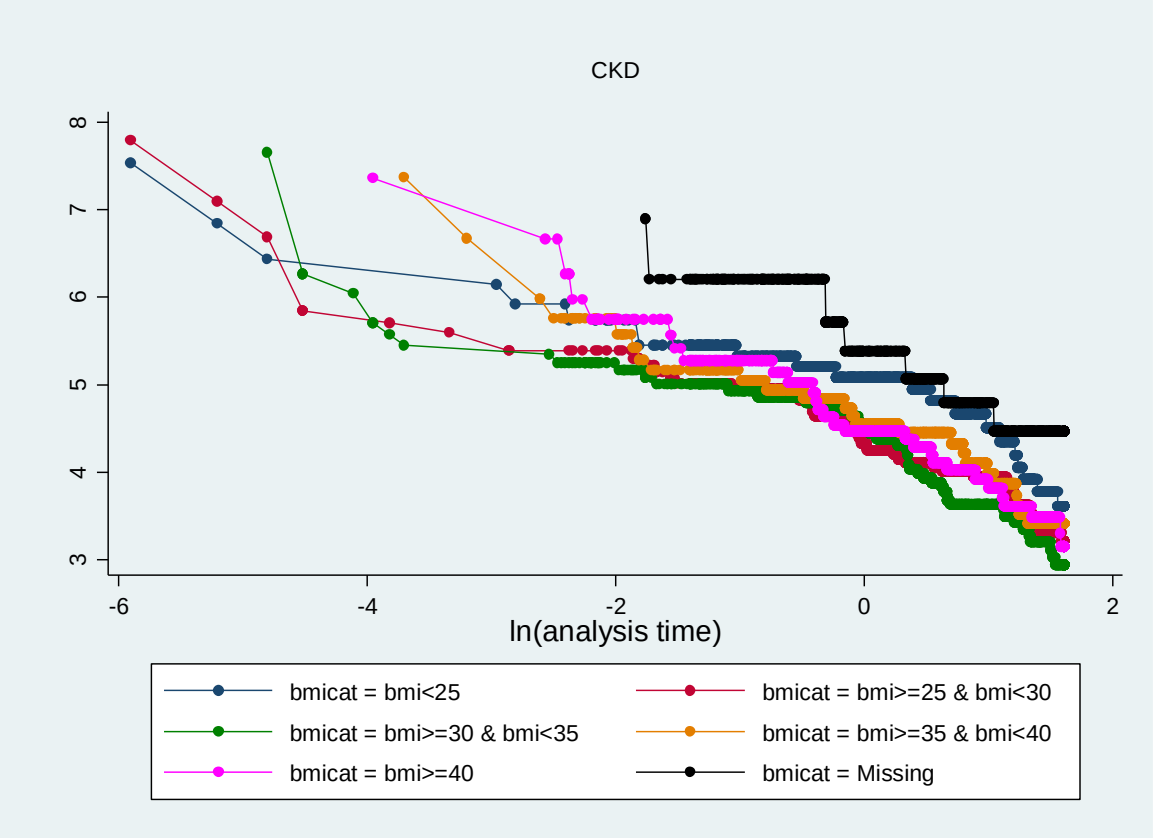
23.0 Appendix14: Adjusted incidence rate of CKD in IGR compared to normoglycaemia

Predictors		IRR	95% (CI)	*P*-Value
IGR vs normoglycaemia		2.6	(2.0, 3.4)	<0.001
Female		1.1	(0.9, 1.4)	0.425
Age		1.1	(1.0, 1.1)	<0.001
Ethnicity				
White	Reference	1.0		
Asian		1.2	(0.8, 1.7)	0.389
Black		1.0	(0.6, 1.9)	0.955
Chinese		1.2	(0.2, 8.7)	0.845
Mixed		0.5	(0.1, 3.3)	0.439
Other		1.0	(0.4, 2.8)	0.970
Missing		0.7	(0.5, 0.8)	0.001
BMI (kgs/m^2)				
<20	Reference	1.0		
20-24.9		1.9	(0.9, 4.5)	0.127
25-29.9		2.5	(1.1, 5.8)	0.029
30-34.9		3.2	(1.4, 7.5)	0.006
35-39.9		2.6	(1.1, 6.2)	0.037
≥40		3.2	(1.3, 7.6)	0.010
Missing		1.4	(0.6, 3.5)	0.452
Deprivation quintile				
Least deprived	Reference	1.0		
2		0.8	(0.5, 1.3)	0.394
3		1.1	(0.8, 1.6)	0.571
4		0.9	(0.6, 1.4)	0.756
Most deprived		1.3	(0.9, 1.9)	0.123
Missing		1.4	(0.8, 2.3)	0.256
CVD		0.9	(0.4, 2.3)	0.905
HF		0.8	(0.1, 6.2)	0.872
AF		2.9	(0.4, 21.0)	0.285
Hypertension		3.1	(2.3, 4.2)	<0.001
NSAID		1.1	(0.9, 1.5)	0.404

Abbreviations: BMI, body mass index; CVD, cardiovascular disease; HF, heart failure; AF, atrial fibrillation; CI, confidence interval

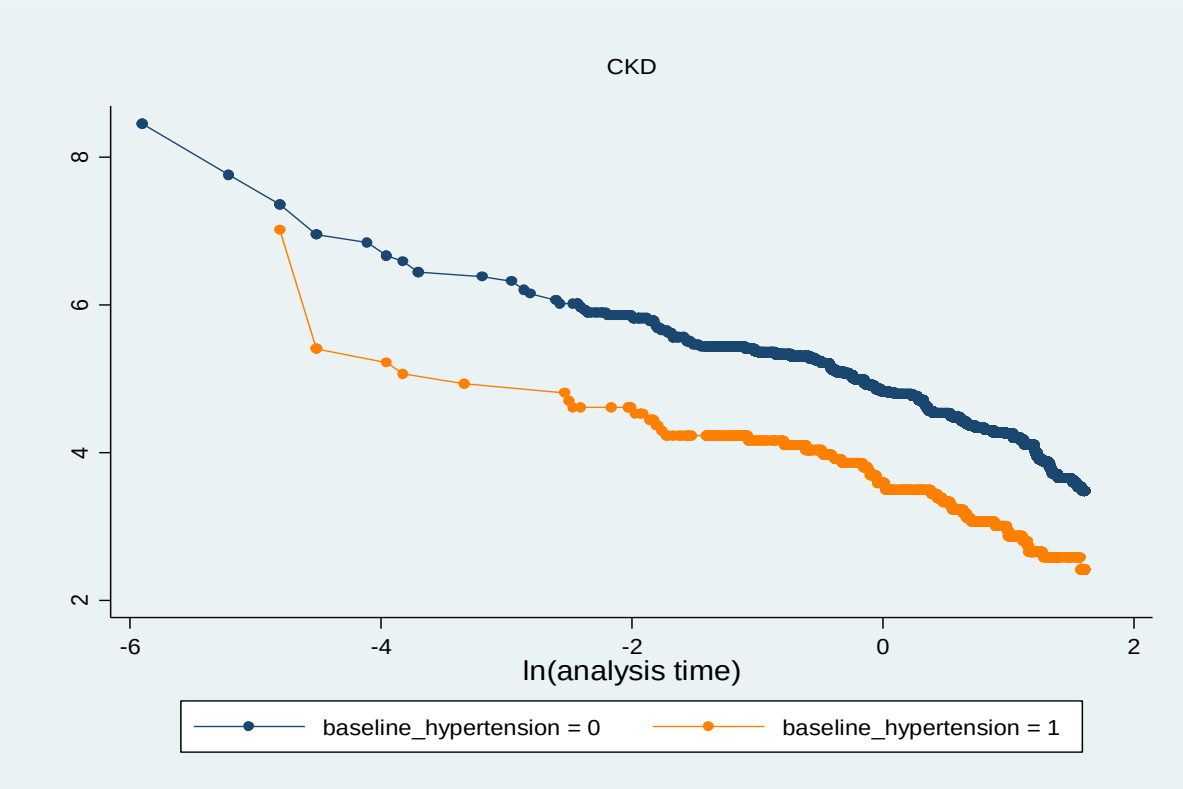
24.0 Appendices 15-17: Log-log plots to check proportional hazards assumption – Five-year follow-up

Appendix15: Log-log plot – BMI categories



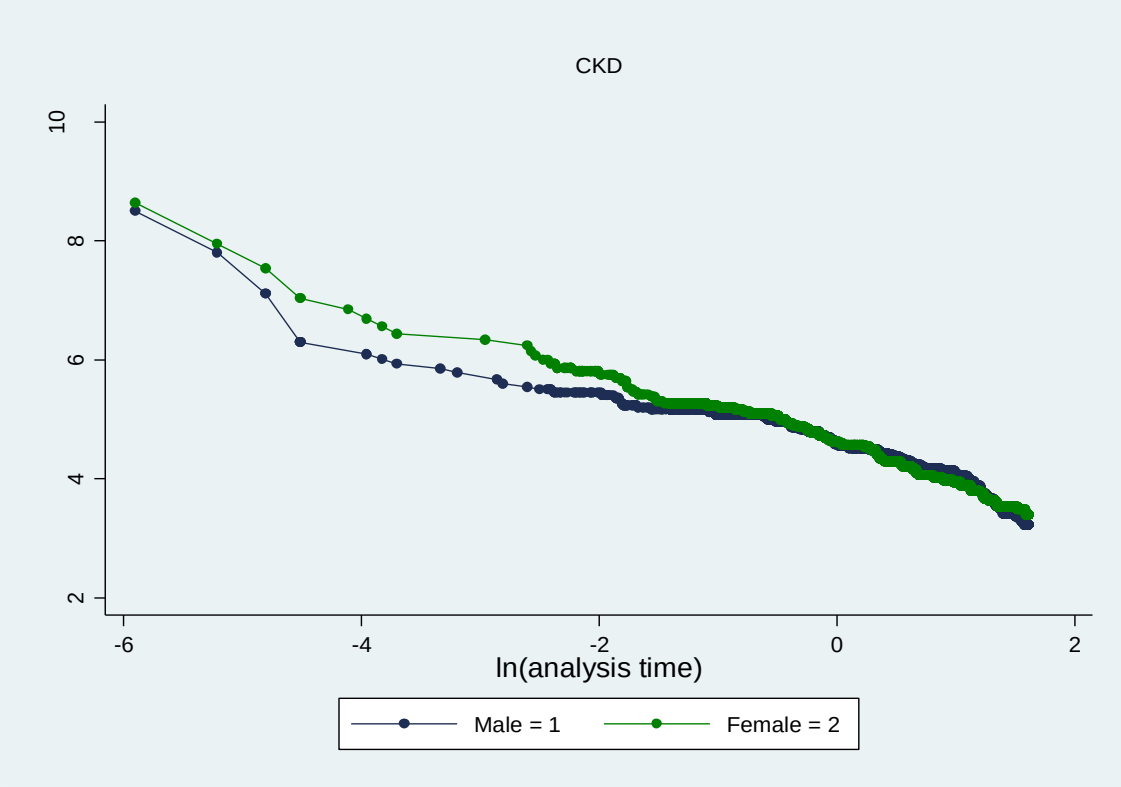
BMI categories: Appendix 15 shows the risk of CKD over time in patients with IGR and various BMI categories. The group remained roughly parallel over time, but the risk of CKD did not increase in a consistent manner for each BMI group. The proportional hazards assumption was not met.

Appendix 16: Log-log plot – Hypertension



Hypertension: Appendix 16 shows the risk of CKD over time among IGR patients who had hypertension to those who did not have hypertension at baseline. The log-log survival curve was parallel meeting the proportional hazard assumption therefore baseline hypertension was included in the model

Appendix 17: Log-log plot – Gender



Sex: Appendix 17 shows the probability of CKD risk over time for males and females. The survival probability remained parallel during follow-up, suggesting that the hazard assumption for both sex was proportional.