# Predicting the epidemic: a study of diabetes risk profiling in a multi-ethnic inner city population
Noble, Douglas James

For additional information about this publication click this link.
http://qmro.qmul.ac.uk/jspui/handle/123456789/2960

QUEEN MARY, UNIVERSITY OF LONDON

**Predicting the epidemic: A study of diabetes risk profiling in a multi-ethnic inner city population**

by

Dr. Douglas James Noble

A THESIS

SUBMITTED TO QUEEN MARY, UNIVERSITY OF LONDON

IN FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTORATE OF MEDICINE

Centre for Primary Care and Public Health

Barts and The London School of Medicine and Dentistry

London, UK

June, 2012

# Abstract

Type 2 diabetes has increased in prevalence globally in recent years, mainly due to obesity. Many other risk factors are well known. Identifying those at high risk of type 2 diabetes may guide targeted interventions aimed at reducing risk.

Type 2 diabetes risk prediction is a complex science. The first half of this thesis presents a quantitative and qualitative systematic review of 145 risk prediction models and scores. Many are available; few are usable in real life clinical practice. Seven have high potential to be used with routine data (such as electronic primary care records).

The second half of this thesis describes the use of one of the risk prediction scores locally, the QDScore, on a dataset of 519,288 electronic primary care records in East London, UK to calculate the ten year risk of developing type 2 diabetes. Ten percent of the population were at high risk (defined as a ten year risk of greater than 20%). Ethnicity and deprivation were key factors responsible for increasing risk, and there was overlap with cardiovascular morbidity. A sub-section of these data were mapped to explore the feasibility of using geospatial mapping to convey the risk of non-communicable disease in a public health setting.

Previous research has focussed on targeting individuals with pre-diabetes (e.g. Impaired Fasting Glucose) and screening for undiagnosed diabetes. Going a step further back and identifying those at risk of type 2 diabetes is theoretically possible due to the wide availability of prediction algorithms, and such an approach is potentially achievable locally using electronic primary care records. This produces important descriptive data

to aid the interventions of general practitioners, public health specialists and urban planners. Future research should focus on interventions which reduce risk of type 2 diabetes in otherwise healthy adults.

# Contents Page

# Tables

# Figures

# Boxes

# Maps

## Acknowledgements

This Doctoral thesis was compiled following a secondment to the Centre for Primary Care and Public Health at Queen Mary, University of London, where I was working on a report on diabetes risk for Tower Hamlets, Newham and City & Hackney Primary Care Trusts. I project managed and led a team working on both a systematic review of diabetes risk models and scores and a cross-sectional analysis of half a million electronic primary care records. There were many others involved in these two projects which form the basis of this thesis.

I would like to especially thank:

Professor Trisha Greenhalgh (primary supervisor) for her supervision of all of my work at Queen Mary, and her advice and help on the systematic review, cross-sectional analysis, geospatial mapping and information governance issues.

Dr. John Robson (secondary supervisor) for suggesting the QDScore as a potential topic for study, for access to electronic primary care data, for support and direction of the cross-sectional analysis, and advice on information governance issues.

Ms. Rohini Mathur for her role as the second reviewer in the systematic review, for extracting and cleaning the data for the cross-sectional study, and assisting with analysis.

The Director of Public Health in Tower Hamlets Primary Care Trust, Dr. Somen Banerjee, and many other general practitioners, clinicians and health service managers in East London, including diabetes leads, who have been immensely helpful during the course of this research in helping me think about the issues at stake, the health information needed, and the possibilities for future interventions.

Professor Sir Liam Donaldson, previous boss and now mentor, for all the skills he taught me, of which the ones on how to handle the media were particularly important for this thesis.

Most importantly, enormous thanks to my wife Alice who supported me to work on this thesis in my 'own time' whilst also completing the final stage of specialist training in public health and preparing for the Harkness Fellowship in the USA.

Finally, the ideas, project management, and written material in this thesis are in majority my own. However, inevitably in a large project like this, where a report and peer reviewed research papers with multiple team members were published before, during, and after the preparation of this thesis, there are figures, boxes, tables, maps, lines and paragraphs of text, and other material, which have been co-designed, co-written and edited with those I have acknowledged above. In addition to the acknowledgments above, full detailed contributorship statements can be found in the relevant parts of the academic papers related to this thesis within Appendices 4, 5, 6, and 7.

## List of Abbreviations

| | |
|---|---|
| ADA | American Diabetes Association |
| AUROC | Area Under Receiver Operating Curve |
| BBC | British Broadcasting Corporation |
| BMI | Body Mass Index |
| BMJ | British Medical Journal |
| CEG | Clinical Effectiveness Group |
| CVD | Cardiovascular Disease |
| EMIS | Egton Medical Information System |
| FINDRISC | Finnish Risk Score |
| FPG | Fasting Plasma Glucose |
| GIS | Geographic Information Systems |
| GP | General Practitioner/Practice |
| HbA1c | Glycated Haemoglobin A1c |
| HDL | High Density Lipoprotein |
| ICD | International Classification of Disease |
| IFG | Impaired Fasting Glucose |
| IGT | Impaired Glucose Tolerance |
| IMD | Index of Multiple Depriviation |
| INEL | Inner North East London |
| MOOSE | Meta-analysis of Observational Studies |
| NHS | National Health Service |
| NICE | National Institute for Health and Clinical Excellence |
| NIGB | National Information Governance Board |

| | |
|---|---|
| NPV | Negative Predictive Value |
| OGTT | Oral Glucose Tolerance Test |
| ONS | Office for National Statistics |
| PCT | Primary Care Trust |
| PPV | Positive Predictive Value |
| QMUL | Queen Mary University London |
| THIN | The Health Improvement Network |
| UK | United Kingdom |

**SECTION 1: Overview**

**Chapter 1: How this research came about**

In December 2010, the opportunity arose as part of my registrar training in public health, to undertake a full-time academic attachment in the Centre for Primary Care and Public Health at Queen Mary, University of London. Previously I had worked in general surgery, as a clinical adviser to the Chief Medical Officer, and as a public health registrar at the Health Protection Agency and Tower Hamlets Primary Care Trust (PCT). My family and I also live in Tower Hamlets. For this reason, I was familiar with the demography of the area, particularly the ethnic diversity and deprivation, and health needs of this part of London.

My professional work-related skills have always been general in nature. My most enjoyable period of time clinically was working in an acting middle-grade position in general surgery in a District General Hospital in Oxfordshire. At the Department of Health I worked on multiple different health policy agendas ranging from paediatrics to patient safety to pandemic flu. At PCT level I enjoyed working across a wide range of commissioning agendas and at the Health Protection Agency I work as first on-call for communicable disease control for a population of 2.7 million.

I met with Professor Trisha Greenhalgh and Dr. John Robson to discuss my academic placement towards the end of 2010. It was clear that there was an interest from all parties in quantifying the risk of developing type 2 diabetes for the local population in Inner North East London (INEL - which includes Tower Hamlets, Newham and City & Hackney – in this thesis referred to as East London). This was also of interest to the

three PCTs in the same geographical area, and this enabled an arrangement whereby my academic attachment would produce a report for the PCTs on type 2 diabetes risk. It was proposed that I undertake a cross-sectional study using an algorithm called the QDScore which had been previously validated to predict ten year risk of developing type 2 diabetes using electronic primary care records. I approached this project enthusiastically as it met the service aim of producing a high quality report for the PCTs that would be useful for commissioning, and also had the potential to meet an academic aim and contribute to the knowledge base and generate a high-quality publication. It also became apparent early on that this could form the basis of a Doctoral thesis.

My first task was to explore the literature of diabetes risk scoring, which I initially found complex. The science of prognostic models and diabetes risk scores was statistically complex, and did not seem to me to be standardised and it was not accessible to a general medical reader. I initially wondered whether it would be possible for me, as a public health doctor interested in the practical application of risk scores in an area of high deprivation and high diabetes prevalence, to continue with the project. As I compiled a set of 29 papers on diabetes risk scores, two things became clear: [a] the heterogeneity between the studies were very high, to the extent that they were barely comparable, and [b] I was surprised that given a 'quick' literature search had produced so many papers, no systematic review existed. Therefore this gave me the idea that perhaps the best way for me to understand this complex literature was to propose doing a systematic review of diabetes risk scores.

Circumstances favoured this choice. The academic centre to which I was attached had an emerging interest in systematic review. Professor Khalid Khan's team had recently

moved from Birmingham and were undertaking Cochrane reviews in women's health, and Dr Catherine Meads from this team was able to offer some input to the diabetes work. Professor Trisha Greenhalgh's team were leading an international collaboration to develop guidance for realist reviews. Some delays had arisen in obtaining the QDScore algorithm and so time was available to work on the review with the funded analyst for the cross-sectional study, Ms. Rohini Mathur. Professor Trisha Greenhalgh was keen to support me conducting the review, and after looking at a few of the papers she suggested that this dataset would be amenable to qualitative (realist) analysis focusing on how (if at all) risk scores were used in real life clinical practice, and what the authors intended to be done with the risk scores they were creating.

We added Dr. Tom Dent to the team, who as a public health consultant had considerable expertise in the applied science of using risk models and scores. From January to June 2011, I project managed and led this team to produce a mixed-method systematic review of diabetes risk models and scores. We reviewed 145 different models and scores and performed an extensive quantitative and qualitative analysis. These were summarised in an academic paper which was published in the British Medical Journal in November 2011.

Although I was correct in assuming that no published systematic review existed when I embarked on this piece of secondary research, the journey took an interesting turn soon afterwards. I learnt that a team from the University of Oxford (where I read medicine) and another team from the University of Cambridge were also conducting systematic reviews on similar but not identical research questions, which they also planned to publish in mid-late 2011. A race against the 'dark blues' and the 'light blues' gave this

work an interesting competitive edge, though the different disciplinary origins of the three principal investigators, public health, statistics and epidemiology, meant that the three reviews, ultimately published within a few months of each other, had a different focus and complemented rather than duplicated one other. Between the three teams we have produced a useful resource on diabetes risk scores from different perspectives. As the National Institute for Health and Clinical Excellence (NICE) were, and are, producing guidelines on risk of type 2 diabetes, all three teams collaborated after the publication of the papers so as to feed consistent advice into that process.

Following submission of the systematic review for publication, the QDScore algorithm became available. Working with Dr. John Robson and Ms. Rohini Mathur it was run on a cross-sectional dataset of 519,288 electronic primary care records, calculating ten-year risk of type 2 diabetes. Multiple pages and tables of data were produced. When I had previously worked at Tower Hamlets PCT, using geospatial mapping to represent Hospital Episode Statistics had been very successful as part of a Care Closer to Home initiative. However, it had not been an academically robust exercise, but had whetted my appetite for the usefulness of geospatially displaying health related data in a public health setting. Professor Trisha Greenhalgh and Dr. John Robson were supportive for me to display the type 2 diabetes risk data in geographical maps. This led to me working closely with Dr. Dianna Smith from the Department of Geography at Queen Mary, who (coincidentally) had conducted her PhD on geographical mapping of diabetes prevalence. Individual postcodes were only available for one of the three PCTs (about a third of the dataset), but with these data from 157,045 records, we mapped risk of type 2 diabetes in Tower Hamlets only. We also decided to explore the use of a new

mapping method called ring mapping and explored the theoretical possibilities of using this for health needs assessment.

In the course of designing the maps we discovered that there was no systematic or consistent information governance guidance for researchers who sought to use postcodes for geospatial mapping. This led to much discussion amongst the research team and with the National Information Governance Board (NIGB) and we explored the issues in detail. There were two principle issues: [a] extracting postcodes (which are considered identifiable data) had to be done in a secure way according to an agreed protocol, and [b] representing small numbers of people on any mapping segment had to be handled in a way which meant it would not be possible to identify an individual.

This second half of the project based on the cross-sectional analysis has resulted in three further academic papers: one on the uses of geospatial maps, which was published by British Medical Journal Open in February 2012; one on the information governance implications of geospatial mapping of small area data which at the time of writing is under review for BioMed Central Public Health; and the results from the overall cross-sectional analysis, which at the time of writing are under review for the British Journal of General Practice. The report to the PCTs contains a condensed version of this Doctoral thesis covering most of the work described above, and some extra data at the level of an individual PCT not included in the thesis.

**Thesis structure**

Chapter 2 concludes section 1 and gives an overview of the epidemiology of type 2 diabetes. Section 2 covers the systematic review of diabetes risk scores. Chapter 3 covers the background, aims and methods of the review, and Chapter 4 the results and discussion.

Section 3 covers the cross-sectional analysis of 519,288 electronic primary care records. Chapter 5 covers the background, aims and methods of the analysis, and Chapter 6 the results and discussion.

Section 4 covers the geospatial mapping of 157,045 electronic primary care records. Chapter 7 covers the background, aims and methods of the analysis, and Chapter 8 the results and discussion. The discussion also includes a consideration of some of the information governance issues and an analysis of the media coverage.

Finally, Section 5 brings together in Chapter 9 the summary of key findings from across Sections 1-3, considers implications for policy, practice and research, and concludes with a personal reflection on the research including lessons learned and key challenges.

I have inserted key tables, boxes, figures and maps in the main body of the thesis to aid readability. The appendices contain copies of the academic papers and relevant checklists and search strategies linked to the systematic review. I have sought throughout to make this accessible to a general medical and public health audience.

# Chapter 2: Type 2 Diabetes Epidemiology

Diabetes is becoming alarmingly common throughout the world.[1] In this chapter I give an overview of type 2 diabetes epidemiology. Whilst working on the systematic review, the cross-sectional study, and the geospatial mapping (all described later), this context provided motivation to quantify and highlight the extent of risk of type 2 diabetes. Although researching risk of type 2 diabetes is different from the epidemiology of established disease, the two are intimately interconnected. Long term efforts to reduce risk of type 2 diabetes has the potential to dramatically decrease incidence. I have divided this chapter into three sections: Global, United Kingdom, and East London. Many different studies and methodologies exist for estimating current and future incidence and prevalence of diabetes. I have sought to summarise headline figures, combined with a more in-depth review of a few selected predictions.

## Global

Almost 350 million people have diabetes worldwide, and the number expected to die from this cause is predicted to double between 2005 and 2030.[2] Type 2 diabetes accounts for approximately 90% of cases. Diabetes is one of four non-communicable diseases – diabetes, cardiovascular disease, chronic respiratory disease and cancers – which account for 60% of global deaths.[3] Their cumulative financial burden worldwide in 2008 was estimated to be US$2.35 trillion and prevalence of disease is projected to increase exponentially.[4] By 2010, the prevalence of diabetes in the adult populations of UK, USA, mainland China and United Arab Emirates had exceeded 7%,[5] 11%,[6] 15%[7]

and 17%[8] respectively.  Americans born in 2000 or later have a lifetime risk of more than one in three of developing diabetes.**[9]**

Global prevalence of both type 1 and type 2 diabetes was estimated for 2000 and 2030, in the year 2000.[10]  Differentiating between type 1 and type 2 diabetes was not possible as this was not coded in most data.  Forty prevalence studies from multiple countries were selected from the literature based on standard criteria for diagnosis of diabetes (e.g. Oral Glucose Tolerance Test – OGTT).  Using these countries' age and sex-specific estimates, extrapolations were made to similar countries without data, based on factors such as ethnicity, socioeconomic status and expert opinion (e.g. Australia's prevalence was applied to New Zealand).   In developing countries, increasing urbanisation was assumed to be linked with double or more the risk of diabetes (e.g. due to reduced physical activity and obesity).  This was as opposed to developed countries where evidence suggested that urban and rural prevalence of diabetes was comparable. Population estimations and projections from the United Nations Population Division were used.  Total number of cases was projected to increase by almost 200 million, from 171 million to 366 million between 2000 and 2030.  This represents an increase in global prevalence from 2.8% to 4.4%.  The Middle East, Africa and India will see the largest relative increases, and the largest total number of cases will be in India (which is predicted to increase from 31.7 to 79.4 million).  More women than men have diabetes, since women live longer and diabetes increases with age, although overall prevalence is higher in men.  These predictions are concerning enough, but are almost certainly an under-estimate since they assume that obesity rates will remain static in developed countries or will follow rates of urbanisation in low income countries.

**United Kingdom**

In the UK (United Kingdom) there are estimated to be 400 new diagnoses of diabetes every day and 90% of cases are classified as type 2 diabetes.[11]  Currently in England approximately 3.1 million people age 16 and above have diabetes.  This figure is expected to increase to 4.6 million by 2030, as a result of ageing, changes in the ethnic composition of the population, and rising obesity prevalence.[5]

An estimation of the prevalence and incidence covering all of the UK between 1996-2005 was published in 2008.[12]  Data on people aged 10-79 years were extracted from the THIN (The Health Improvement Network) database, a collection of almost 5 million electronic general practitioner records.  After exclusion criteria were applied 1.84 million records were analysed.  Incident cases were defined as those recorded as a new diagnosis between 1996-2005 (numerator - number of new cases/denominator - person-years of patients at risk).  Prevalence was defined as the number of existing cases when the study commenced plus the incident cases (numerator – new and old cases/denominator – total number of records).  During this period of time there were 41,386 incident cases of type 2 diabetes.  Prevalence of type 2 diabetes increased from 2.47% to 3.9% between 1996-2005.  Incidence of type 2 diabetes increased from 2.60 to 4.31 per 1000 person-years between 1996 and 2005 (age and sex standardised).  The change in incidence was less in men (63% increase) than women (69% increase).

Obesity associated with incident cases of type 2 diabetes rose from 46% in 1996 to 56% in 2005.  The researchers suspected much of the increase in type 2 diabetes was therefore due to obesity.  This finding is particularly concerning when considered with

the global estimates described previously,[10] which leads to the likely conclusion that the projected worldwide prevalence in 2030 will indeed be much higher than 366 million.

In the course of this research both ethnicity and deprivation emerged early on as key determinants for increasing risk of type 2 diabetes. This was consistent with existing evidence on those with established disease. For those diagnosed with type 2 diabetes wide social and ethnic differences in prevalence are well recognised. Deprived populations have higher rates of type 2 diabetes[13] likely linked to higher rates of obesity and lower income in these population groups. South Asians and Black groups have a rate of type 2 diabetes six and three times greater than the White population respectively[13] and the most affluent fifth of the whole population (<55 years) have half the prevalence of the most deprived fifth.[14] The reasons for ethnic associations are complex and not fully understood, but evidence suggests South Asians may be more prone to fat deposition patterns that predispose to type 2 diabetes[15-17], and ethnic minorities in general are less likely to exercise.[13]

The burden of type 2 diabetes related morbidity and mortality is considerable for both patients and local and national health economies. Expenditure on diabetes in the National Health Service (NHS) may be as high as 10% of total yearly budget[18] and 10-20% of patients in hospital have diabetes (this cohort also has disproportionately longer in-patient episodes and increased costs).[11] As many as 50% of people with type 2 diabetes have complications at diagnosis,[19] which may have been detectable up to seven years previously, and the onset of type 2 diabetes as long as twelve years prior to formal diagnosis.[20]

**East London**

The cross-sectional analysis and geospatial mapping described later in this thesis used data, as stated before, from three inner city boroughs and one city district (which is grouped with one of the boroughs): Tower Hamlets, Newham, and City & Hackney. Using data from the Quality Outcomes Framework (a remuneration programme for all general practitioners (GPs) based on clinical performance) overall numbers of registered persons over 16 with diabetes has been estimated to be 45,688, which is approximately 6% of the population (Tower Hamlets 13,770; Newham 18,467; City & Hackney 13,451).[21] Quantifying prevalence of type 2 diabetes is problematic due to undiagnosed disease, uncertain population sizes, high population mobility and inaccuracies in disease recording and registration data. Local data from the Clinical Effectiveness Group (CEG) estimate a standardised prevalence of approximately 7% (10% in the South Asian ethnic group). Diabetes registers have been steadily increasing. In 2010 the register size was 40,866. These changes are more marked in Tower Hamlets and Newham.[22]

The Borough of Newham has the highest prevalence of diabetes of all boroughs in London. Increasing rates in children and young people are thought to be linked to the increase in obesity (Newham has the second highest rates of obesity for reception year children in England). The non-white ethnic groups are at increased risk. Doctor-diagnosed diabetes was 2.5-5 times greater than the general population for Black Caribbean and South Asian groups. It is common in Newham for diabetes to be thought of as a normal part of life, and it is of concern that this may have affected the population's view of the serious health consequences.[23]

The City of London & the Borough of Hackney has the lowest prevalence of diabetes in East London. Yet, the standardised rate of registered persons with diabetes on the register has increased from 4.2% - 5.4% between 2004-2010, compared to 5.9% - 7.4% for Newham and 5.9% - 7.2% for Tower Hamlets.[22] It is also recognized locally that there is a large undiagnosed population. In the over 40 age group prevalence is 11.1%, and GP level prevalence in this group vary from 4.7% to 20.3%.[24]

Late diagnosis and co-morbidities feature strongly as characteristics of the existing diabetic population in Tower Hamlets. In 2006, Picker Institute Europe surveyed people with diabetes from ten general practices.[25] Out of 856 people selected at random from the practices, 340 people returned completed surveys. The poor response rate means caution needs to be taken with the findings. Eleven percent reported type 1 diabetes, 53% type 2. Findings indicated that for some people diagnosis was made late – 16% started insulin within three months of diagnosis. The majority (40%) were middle aged (40-59 years), and 77% were on medications for other conditions. High prevalence overall in Tower Hamlets is partly attributed to the increased risk of disease in South Asian communities (who make up 30% of the population). Large increases are expected overall, with an average prevalence of 10.1% expected by 2030.[26]

**Conclusion**

Diabetes is increasing exponentially: globally, in the UK, and locally in East London. Estimates are fraught with various difficulties, including: [a] coding being unable to distinguish between different types of diabetes, [b] under-estimations due to other

unaccounted for factors, [c] people with undiagnosed disease not being accurately accounted for (known unknowns), [d] uncertain future population sizes, and [e] poor response rates in survey data potentially over or underestimating findings. Yet, increases in prevalence and incidence are undisputed and the contribution of obesity, lack of physical activity, ethnicity and deprivation are known to be major risk factors.

These findings from the research literature offer strong support for initiatives aimed at preventing diabetes. In the next section, I will describe how I approached this by reviewing scoring systems designed to identify those at high risk of developing type 2 diabetes in the future.

## SECTION 2: Systematic Review of Diabetes Risk Scores

## Chapter 3: Background, Aims and Methods

This section presents a systematic review of type 2 diabetes risk models and scores. The background section of this chapter describes the history of diabetes risk scores, and the underlying science. I then outline the aims of the systematic review including detailed methodology, and changes that occurred during the course of the study.

**Background**

Risk prediction is a complex science concerned with estimating the likelihood of any individual (or population) developing an outcome (usually a disease) of interest, based on a series of known risk factors. Epidemiologists and statisticians have been striving to produce weighted models and scores that are perceived as sufficiently simple, plausible, affordable and implementable to be adopted widely in clinical practice.[27][28]

Risk models and scores first emerged prominently in relation to cardiovascular disease, and these are widely used in clinical and public health practice. In the UK for example, all general practice electronic patient record systems offer the facility to calculate the 'Framingham score' or 'QRisk score', a patient's 10-year risk of a cardiovascular event. These cardiovascular disease risk scores feature in many guidelines and decision pathways (such as the cut-off for statin therapy[29]) and general practitioners receive financial rewards for calculating it.[30]

Researchers in England, America and Finland have led the development of type 2 diabetes risk models and scores. Presented below is a brief description of three early models and scores, one from each of these countries, to set the scene for a more detailed discussion of the individual steps in creating a risk score.

The Cambridge Score[31] is well known and widely cited in the literature. Developed by Griffin et al. in 2000, it used cross-sectional data to derive a risk score. These researchers determined that age, gender, body mass index, steroid and antihypertensive medication, family and smoking history were most predictive of risk in an English population aged 40-64 years. This was also one of the first studies that showed that information routinely held on GP databases could predict type 2 diabetes. The risk score was developed retrospectively from two separate cross-sectional research populations. First, from general practices in Cambridge (Ely study, population size 4,922). And, second, from a population of registered people from 41 general practices in Wessex (population size 455,566). In total 1077 randomly selected people from the Ely study without diabetes aged 40-64 completed a standard OGTT, lifestyle questionnaire and clinical measurements. All new diagnoses of type 2 diabetes (48) were recorded. The Ely study group was randomly divided in two, one half to derive the risk score and the other half to internally validate (see definition in Table 1) the score. To derive the risk score half the group of newly diagnosed type 2 diabetics in age groups 40-64 years (25 cases) were added to half of the newly diagnosed type 2 diabetics in age groups 40-64 years (101 cases) from the Wessex group. This formed a larger group for deriving the score and their variables were used in a logistic regression with risk factors being retained if statistically significant ($p<0.05$). This study illustrated the complex methodology and study designs which would become common

in this field, and the Cambridge Score has heavily influenced future researchers deriving and validating risk models and scores.

The San Antonio Score[32] created in 2002 was one of the first scores to incorporate biochemical indices.  Medical history, Body Mass Index (BMI), blood pressure, OGTT, fasting serum cholesterol and triglyceride levels were obtained retrospectively from the San Antonio Heart Study.  Random selection from this cohort resulted in a population of 1,791 Mexican Americans and 1,112 non-Hispanic Whites without known type 2 diabetes from a possible population of 5,158 (which was based on randomly selected households representing different socio-economic and ethnic groups, age range 25-64years).  A full model and a simplified model were developed and comparisons between scores for identifying type 2 diabetes performed - both including and excluding the OGTT.  The cohort was followed for 7.5 years and 204 Mexican Americans and 65 non-Hispanic Whites developed type 2 diabetes.  The clinical model without OGTT consisted of: age, sex, ethnicity, fasting glucose level, systolic blood pressure, high density lipoprotein (HDL) cholesterol level, BMI and family history of diabetes.  This study was one of the first to use many different combinations of variables and compare their statistical properties.

The Finnish Risk Score[33] (known as FINDRISC) published in 2003 is currently the most cited diabetes risk score.   It used the National Population Register in Finland in 1987 to create a score based on a random sample of 4,435 people (aged 35-64 years) without a history of using diabetes medication.  Individuals were followed for ten years to determine who would develop drug treated diabetes by using the Social Insurance Institution drug register up to 1997.  182 people developed incident type 2 diabetes.  A

second sample was taken from the National Population Register from 1992 and used as a validation cohort. The final score consisted of: age, BMI, waist circumference, blood pressure medication, history of high blood glucose, physical activity and various foods eaten. The Finnish risk score set various precedents, including making use of wide ranging data-sets, making assumptions and using proxies for diagnosis of type 2 diabetes, and toggling variables to develop 'full' and 'concise' models.

These three early risk scores were important in the development of this field as they established a number of methodological approaches as 'standard' in the literature, which were to heavily influence future researchers, including: [a] complex study designs which were unlikely to be easily repeatable by other researchers, although principles could be applied, [b] testing various models with different variables, [c] innovatively using various data sources to identify people without diabetes, and [d] using different methods for confirming diagnosis of type 2 diabetes.

This brief review of three early models demonstrates that diabetes risk scores, which predict risk of developing future disease, can broadly be considered as prognostic models and should broadly be subject to the same methodological principles[34-37], although the key difference is they are considering risk of future disease as opposed to outcomes of existing pathology. From an initial literature review of 29 papers and discussion with various experts in the field, it became clear that there were several steps involved in creating, testing and using a type 2 diabetes risk score. These six steps are shown in Table 1, and then discussed in detail as part of the context for the formal systematic review.

| Six steps of type 2 diabetes risk score development | Definition |
| --- | --- |
| Derivation | Creating a diabetes risk score composed of various individual risk factor components |
| Discrimination | Distinguishes reliably between people who will develop the condition and people who will not |
| Calibration | Assessing over time whether what is initially predicted transpires to be close to what is actually observed |
| Validation | Running the score on a different population of people to test its performance |
| Reclassification | Comparing different risk scores with different variables around given thresholds |
| Application | Using the score in real life on an actual population e.g. in a general practice surgery |

**Table 1: Development of a diabetes risk score**

The gold standard approach for deriving a type 2 diabetes risk score is to take a large, age-defined, non-diabetic population cohort, measure baseline risk factors, and follow the cohort for a sufficiently long time period to see which individuals go on to develop diabetes.[37] An example of a cohort used for this purpose is the Framingham heart study, which started in 1948, and is a well known example of a cohort study that followed a population over several generations to study risk factors for cardiovascular disease (CVD).[38] Assessing risk of type 2 diabetes from existing research populations, like Framingham, or from routinely collected data, has been particularly attractive as this approach uses readily obtainable information without the need for de novo data collection or for invasive procedures such as blood glucose testing.

Re-testing for type 2 diabetes at the end of the cohort should ideally use the same test as at baseline. Follow-up rate should be as high as possible. The final sample therefore has two principle groups: those with and without type 2 diabetes, all with a series of measured risk factors at baseline. This provides a basic incidence statistic over time. A

regression analysis of risk factors determines which were statistically correlated with disease, allowing the creation of a weighted model.

An example of an approach to score derivation can be worked through by considering a risk score devised in 2006 in Thailand based on a 12 year cohort of people from 1985 and 1997.[39] The initial cohort comprised 3,499 workers aged over 35 years from an electric plant. The initial cohort had been created prospectively to study vascular risk factors contributing to cardiovascular mortality. Multiple baseline measurements were collected including a fasting plasma glucose (FPG) and OGTT. This allowed exclusion of people with type 2 diabetes based on the American Diabetes Association (ADA) definition of type 2 diabetes from the FPG and OGTT results. An additional exclusion criteria of a previous known diagnosis of diabetes was also used. Within the baseline measurements were multiple risk factors for developing type 2 diabetes. The researchers chose to analyse: sex, age, BMI, waist circumference, hypertension, family history, smoking, alcohol consumption, IFG, impaired glucose tolerance (IGT) and serum cholesterol. The cohort began with 3,254 people and 2,667 (82%) were present at the end of the study. At the 12 year point 361 had been diagnosed with type 2 diabetes defined by: FPG, OGTT, prescription of diabetes medication or diagnosis. Incidence of diabetes was calculated as 13.5% in this cohort over 12 years. The differences in the values of the risk factors between the 361 with type 2 diabetes and the 2,316 without type 2 diabetes were compared using a logistic regression with diagnosis of type 2 diabetes as the dependent variable and each of the risk factors as independent variables. P-values were displayed for each risk factor indicating whether the difference was statistically significant. Odds ratios for predicting diagnosis of type 2 diabetes at 12 years with 95% confidence intervals for each risk factor measured at inception was

calculated.  Certain risk factors were significant and these were included in the final score.  The log of the odds ratios from these factors was presented as a beta-coefficient and a numerical score as an integer attached to each.  The final score was out of 17 as shown in Table 2.  A score of 17 equated with a 100% cumulative incidence of diabetes at year 12.

| Risk factor | Coefficient | Diabetes risk score |
|---|---|---|
| **Age (years)** | | |
| 34–39 | | 0 |
| 40–44 | -0.07 | 0 |
| 45–49 | 0.27 | 1 |
| ≥50 | 0.60 | 2 |
| **Sex** | | |
| Women | | 0 |
| Men | 0.44 | 2 |
| **BMI (kg/m$^2$)** | | |
| <23 0 | | 0 |
| ≥23 but <27.5 | 0.69 | 3 |
| ≥27.5 | 1.24 | 5 |
| **Waist circumference (cm)** | | |
| <90 in men, <80 women | | 0 |
| ≥90 in men,  ≥80 in women | 0.56 | 2 |
| **Hypertension** | | |
| No | | 0 |
| Yes | 0.64 | 2 |
| **History of diabetes in parent or sibling** | | |
| No | | 0 |
| Yes | 1.08 | 4 |

**Table 2: Example of creation of diabetes risk score**

*Reproduced from Aekplakorn W, Bunnag P, Woodward M et al. A risk score for predicting incident diabetes in the Thai population. Diabetes Care 2006;29(8):1872-77.*

The coefficients in Table 2 above represent what I have described broadly as the underlying *model*, which in some studies is presented as a more complex mathematical equation.  This is linked, but separate from the risk *score*, which is generally presented as an integer derived from the coefficient.  I have often used the phrase 'type 2 diabetes

risk model or score' reflecting the sequence of derivation events. A risk score based on regression analysis has an underlying model, but not all models have been converted to risk scores.

Current cohort derived type 2 diabetes risk models and scores have generally been derived from retrospective research on population cohorts which were assembled for another purpose. This includes risk models and scores which claim to have been derived from 'prospective' cohorts. The term 'prospective' is accurate insofar as the inception cohort was followed prospectively (for example to look at the development of some other disease such as cardiovascular events), but it is incorrect insofar as the study of *diabetes* risk is strictly retrospective. The 'prospective' cohort is typically studied at a future point in time (e.g. a diabetes risk researcher in 2010 may begin to study a cohort assembled and studied between 1990-2000), thus achieving 'retrospective research of a prospective cohort'. This opens up the potential for a series of biases, including: [a] the initial population has a degree of selection bias (the cohort was assembled for a different research purpose), [b] lack of sensitivity to diagnose type 2 diabetes as cohorts may have been designed to primarily detect different outcomes, [c] being restricted to a set of pre-defined risk factors selected by different researchers for a different purpose, and [d] the instruments used to measure risk factors (such as lifestyle questionnaires) may not have been specifically validated for type 2 diabetes, introducing a possible measurement bias. In addition to possible biases, other limitations are inherent. For example, specific age ranges are often used for cohort studies (e.g. 35-65 years) and those using the cohort for type 2 diabetes risk score generation later are therefore limited to this age range.

Considerable variability also exists in methodologies employed to derive a diabetes risk model or score from a cohort study. The original research team will typically have recruited participants at a start date or over a fixed period of time. To derive the risk model or score for type 2 diabetes a series of exclusion criteria are usually applied to entrants, most importantly removing all persons with known diabetes. Other exclusion categories include pregnant women (since development of gestational diabetes would influence results), and children and adolescents (since type 2 diabetes is less common in these groups). Baseline data are often collected at cohort inception. These include variables such as age, gender, ethnicity, medication history and BMI. Tests for existing diabetes, such as OGTT or diabetes medication prescriptions are sometimes used to exclude participants with diabetes and confirm that undiagnosed diabetes is not present within the cohort. The cohort of individuals is followed for a period of time, typically between 3-15 years. Re-measurement of some or all of the baseline variables takes place at points throughout the study and/or at the end of the study period. At the intervals and end point of the study a number of participants will have developed diabetes, a basic incidence statistic. As with the baseline tests, the diagnosis of diabetes may be confirmed in a variety of ways between different studies. Box 1 below shows some of the ways that a diagnosis of type 2 diabetes is assumed at inception and at various points throughout cohort studies.

```
FPG
OGTT
Death certificates
Physician diagnosis
Self-reported diagnosis
Medication registers
HbA1c level
International Classification of Disease (ICD) code
Dietary treatment
Medical records
Electronic records
```

**Box 1: Sources of diagnostic criteria for diabetes in risk score studies**

Multiple diagnostic criteria present numerous problems, including inaccurate diagnoses and lack of comparability between studies. Additionally, many of these criteria rely on coding which often has high levels of associated errors, and using a different diagnostic test at the beginning and end of a cohort may produce inconsistent prevalence estimates and threaten both accuracy and precision.

Once a diabetes risk score has been derived, an assessment of how reliably it predicts those who go on to develop diabetes needs to be made. One of the most common measures used is discrimination. This is often measured as sensitivity and specificity, or a function of both. High sensitivity in this context refers to the likelihood that the risk score predicts who will develop diabetes. For example, a test with 75% sensitivity correctly identifies 75% of those who will develop diabetes, and misses 25%, incorrectly classifying them as negatives. The term 'specific' in this context means that the risk score reliably picks up those who will not go on to develop diabetes e.g. a test with 75% specificity correctly identifies 75% of the persons who will not develop diabetes, and misses 25%, incorrectly classifying them as positives.

Sensitivity and specificity of a risk score can be combined to produce an Area under a Receiver Operating Curve (AUROC). This is a graphical display of these two measures and helps visualise discriminatory ability. On the y-axis sensitivity is plotted versus 1-specificity on the x-axis. 1-specificity is used as it measures the rate of false positives (those who will not get the disease but are falsely classified as positives by the risk model or score using that threshold). Sensitivity is an indication of the true positive rate, and so the graphical display allows true positives and false positives to be compared. This creates a curve showing the change in 1-specificity (or false positive rate) for increasing sensitivity (or true positive rate). There is a trade-off between sensitivity and specificity and usually the optimum cut-off is the point of the highest true positive and lowest false positive rates. The AUROC can be quantified, known as the C-index, a statistic by which different risk models and their discriminatory power may be compared.

A cut-off score is usually determined which picks up the highest number of true positives and the lowest number of false positives. This is unlikely to be the highest score, although it could be. It may be that a score of e.g. 17/20 has a sensitivity of 75% and a specificity of 65%; however a score of 12/20 has a sensitivity of 95%, but specificity of 40%. Research studies often tabulate these results and show the cut-off score that has the highest sensitivity and specificity, or represents it as the AUROC.

If significant risk factors are missed out of a score (as may happen systematically during 'retrospective research of a prospective cohort'), the score will produce a consistently biased result – either under or over estimating low or high risk. The example of ethnicity as a known significant risk factor for type 2 diabetes highlights this problem.

In a population of entirely White ethnicity the variables age, gender, BMI, waist circumference, hypertension, and history of diabetes in parent or sibling will predict risk uninfluenced by ethnicity as a risk factor. The same would be true in an entirely South Asian population, because relative to each other the risk is the same, but overall the risk would be underestimated as it did not factor in the increased risk due to South Asian ethnicity. However, this would not be reflected in the discrimination statistics. In practice, many populations are heterogeneous with respect to ethnicity. For example, in a population with 20% South Asian ethnicity and 80% White ethnicity the score would perform adequately for the White population but would underestimate the risk for the South Asian population. As ethnicity would be unknown, as it was not a measured variable, this would mean a poorer result for those truly at high risk of developing type 2 diabetes i.e. some South Asian people would falsely be classified at lower risk when they were in fact at higher risk. In this example the larger the South Asian group the more inaccurate the result at a population level. The reason for this would not become clear until ethnicity was added as a variable.

Calibration is an essential statistic to accompany discrimination and is a further measure of predictive power. It assesses whether what is initially predicted transpires to be close to what is actually observed over time. This is often presented as an observed to predicted ratio or a statistic which describes this such as the Hosmer–Lemeshow test or Brier score. By definition this has to occur over a real period of time, and this period of time has to be clinically meaningful. Therefore, an observed to predicted ratio of disease occurrence for type 2 diabetes should only be applied to a study which covers a period of time that will allow type 2 diabetes to start occurring in reasonable numbers within the population in question. This is also dependent on age structure; older

populations being more likely to have higher disease occurrence rates over any period of time. There are no hard and fast rules to govern study length, but given that diabetics should have been excluded via diagnostic testing at inception, it is not unreasonable to suggest that four years should be a minimum study period (based on type 2 diabetes complications being detectable 4-7 years prior to diagnosis[20]). By definition only longitudinal cohort derived risk models and scores can be calibrated.

Validating a risk score refers to using the score on a different population of people to test its performance, usually in terms of discrimination and calibration, but this also offers an opportunity to assess usability in a different context. Validation can be done either internally (by splitting the original sample, developing the score on one part and testing it on another), temporally (re-running the score on the same or a similar sample after a time period), or externally (running the score on a new population with similar but not identical characteristics from the one on which it was developed).[34 37] External validation on a separate cohort with demographic differences to the original population is the gold standard to assess whether a model can be used accurately outside of the population it was developed on.

Comparison of discrimination between different models (especially those that have added a 'new' variable) is common in research deriving and validating type 2 diabetes risk scores. To markedly improve already adequate discrimination in a model of basic risk factors (such as age, sex, and ethnicity) 'new' variables have to demonstrate independent and significant links to risk of type 2 diabetes.[40] Yet, this alone does not confer a better model in clinical practice, particularly where a threshold may exist for a change in clinical management. More important is the performance of the model at

different thresholds.[41]  For example, locally in East London if an individual has a greater than 20% risk of cardiovascular disease using QRisk[42] during the NHS Health Checks programme[43] in general practice, this triggers further clinical management.  At a population level the proportion who fall just above and below 20% risk become significant, especially if the intervention (or lack of) is costly e.g. an invasive investigation or prescribing medication, or missing those at high risk with consequent long term prognostic implications.  This problem has given rise to the concept of reclassification (performance around a pre-defined threshold[44]) and accompanying statistical tests, including the reclassification calibration statistic.[45]  This is derived from the Hosmer–Lemeshow test and contrasts observed to expected results at various thresholds.  Different risk scores with different variables can be compared using this technique.   Net reclassification improvement compares risk reclassification and tabulates positive and negative change around given thresholds.[45]  This can reveal large changes in who falls above and below the threshold for clinical management, even if the change in discrimination overall is small.[40]

Caution should be exercised when extrapolating a risk model or score developed in one population or setting to a different one (e.g. secondary to primary care, adults to children or one ethnic group to another).[36]  Using a score which has been derived from a demographically different population will likely result in poorer performance in a new population, unless the majority of the most significant risk factors for that population (e.g. ethnicity and deprivation) are included in the score.  Ideally there would be a unique risk score for each population.  However, in most cases this is impractical and expensive, and so a best-fit model needs to be found, based on the performance tests described above, and an assessment of usability.  On this latter point this type of

qualitative assessment has been done infrequently, and therefore became a major focus of the systematic review.

**Cohort versus cross-sectional designs for diabetes risk scores**

In the course of writing this thesis and carrying out the systematic review, I encountered a different, but commonly used approach to develop a type 2 diabetes risk model or score by using data from a cross-sectional survey. Researchers applying this approach have generally started with a population at one point in time who are tested for type 2 diabetes and have certain clinical variables measured. This splits the population into those with and without current disease. A regression analysis is applied, clinical variables weighted, an equation created, a score derived, and performance assessed.

This approach has several problems. Those with type 2 diabetes have not been studied from the point of being healthy subjects to developing disease. The measurement of clinical variables are not truly risk factors, they are instead disease characteristics. This makes temporal association impossible i.e. is the measured variable a risk for developing disease or is it a consequence of pathology – under or over estimation of the significance of any given disease characteristic is likely. Discrimination describes the ability of the model to detect current disease and calibration is not possible as there is no observed to expected ratio. Figure 1 pictorially represents the main differences.
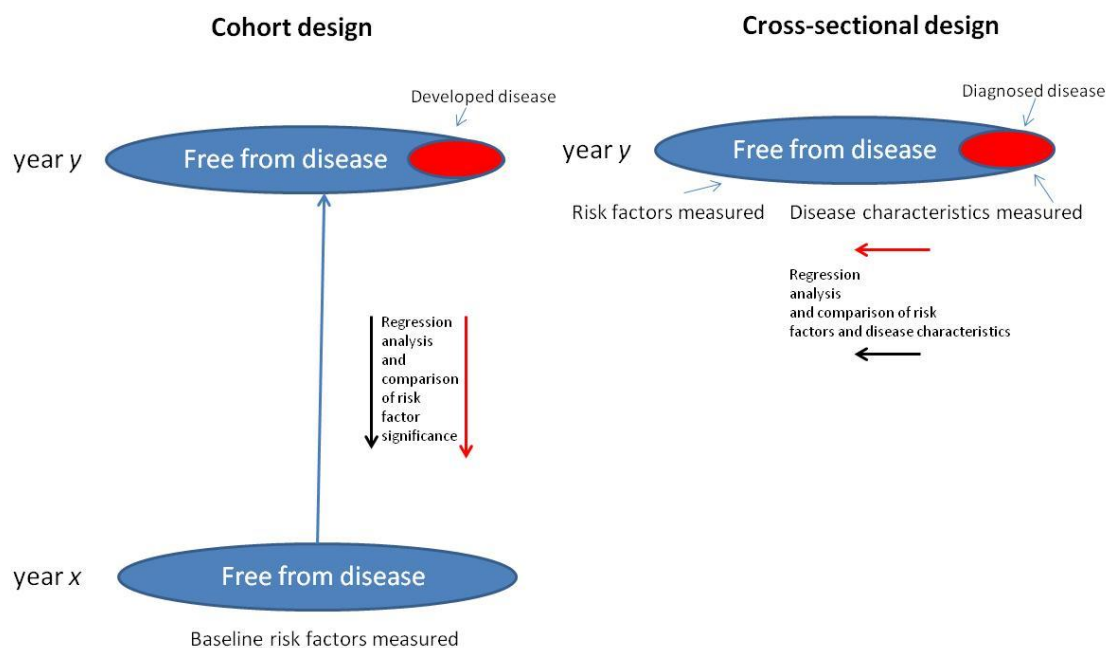
**Figure 1: Cohort versus cross-sectional design for deriving a risk score**

A further problem is the conflation of incidence and prevalence. Incidence is defined as the number of new cases of disease over a specified period of time. Prevalence refers to the total number of cases in any one population at a given point in time. Using prevalence data to predict incidence is only valid in very limited circumstances. For example, in acute infectious disease outbreaks of short duration (e.g. Norovirus) prevalence and incidence are closely correlated as the disease in question has low background rates in any one population and has a short duration of infection. However, for chronic diseases, using prevalence data (the cross-sectional approach) to estimate incidence is problematic.

Though prospective longitudinal designs in specially assembled cohorts are expensive, difficult and time-consuming to execute, cross-sectional designs in which risk factors are measured in a population including both diabetic and non-diabetic individuals are methodologically inferior.

**Implications of the literature for further research**

It is often proposed that risk scores and other prognostic models should be subject to 'impact studies' – that is, studies of the extent to which the score is actually used and leads to improved outcomes. Whilst most authors emphasise quantitative evaluation of impact e.g. via cluster randomised controlled trials,[36] much might also be learnt from qualitative studies of the process of using the risk score, either alone or as an adjunct to experimental trials. One such methodology is realist evaluation, which considers the interplay between context, mechanism (how the intervention is perceived and taken up by practitioners) and outcome.[46] In practice, however, neither quantitative nor qualitative studies of impact are common in the assessment of risk models and scores.[36]

It has been suggested elsewhere that appraisal of risk prediction models have three core areas.[41 44] First, the *context* for use of the risk prediction model, including the disease of interest, the population that the model will be applied to, and the evidence base that changing risk is possible. Second, the actual appraisal of the *performance* of the model, including discrimination, calibration, reclassification and validation, and an assessment of the quality of the underlying data. And, third, *implementation* in real life, including costs, ethics, and training and prioritisation considerations.[41]

Background reading and a basic literature review revealed that whilst there were multiple models and scores for assessing risk of developing type 2 diabetes, none were in routine use in the UK, in general practice or public health. There was also no systematic review in the academic literature, despite many different models being

available. I thought this presented a confusing picture for both GPs and public health specialists, who would be potentially faced with a very complex literature, multiple different methodologies, and probably very few studies of use in real life. In view of this situation I wanted to comprehensively determine the performance and impact of risk models and scores for predicting type 2 diabetes in adults.

**Aims of the systematic review**

In this systematic review of type 2 diabetes risk models and scores the aim was to inform the selection and implementation of diabetes risk models and scores.

The research question for this component of the thesis was:

*What is the performance and impact of risk models and scores for predicting type 2 diabetes in adults?*

Hence, I was particularly interested in highlighting the characteristics of a risk model or score which would (if appropriate) increase its adoption and use in practice. To that end, I sought along with other members of the review team, comprising Greenhalgh, Dent, Meads and Mathur to review the literature on development, validation and use of diabetes risk models and scores in different contexts and settings, using [a] quantitative data on demographics of populations and statistical properties of models and scores, and [b] qualitative data on how models and scores were perceived and used by practitioners, policymakers and others in a range of contexts and systems. We decided given the methodological problems with cross-sectional models and scores, as described previously, to base the systematic review on cohort studies only, although this decision was not taken initially and evolved as the review progressed.

**Objectives**

1. To systematically review known type 2 diabetes risk models and scores for adults.

2. To analyse the demography of the populations from which the models and scores were derived.

3. To analyse the demography of the populations on which they were validated.

4. To analyse the final components of the models and scores and their contribution to overall risk.

5. To compare the discrimination and calibration statistics used for quantifying risk.

6. If possible and valid, to perform a meta-analysis with specialist statistical support.

7. To review qualitatively the purpose of the models and scores and their use in clinical and public health practice.

**Methods**

Below, I describe the methodology and study protocol. A tabulated version of the study protocol and MOOSE (meta-analysis of observational studies) checklist were submitted to the British Medical Journal (BMJ) along with the manuscript of the systematic review. These two documents can be found in Appendix 1 and 2 respectively.

A scoping search was undertaken in January 2011, focusing mainly on existing well known type 2 diabetes models and scores recommended by experts and contextual

background material. The yield of 29 papers from this search was used to develop the protocol for the review, including search terms, and inclusion and exclusion criteria.

An information scientist, Helen Elwell (HE), a librarian at the British Medical Association Library, helped design a search strategy. She was assisted by Catherine Meads (CM) and myself (DN, principal investigator). Relevant guidance in *Systematic Reviews: Centre for Reviews and Dissemination guidance for undertaking reviews in health care*, and *Systematic Reviews to Support Evidence-Based Medicine* was drawn on to identify any relevant studies of type 2 diabetes risk models and scores.[47] [48]  I implemented the final search strategy, which was double-checked by both HE and CM. The final search was undertaken on 11th February 2011.

The key words for the literature search were: predict, screen, risk, score, [type 2] diabetes, model, regression, risk assessment, risk factor, calculator, analysis, sensitivity and specificity, ROC and odds ratio. A decision was made to search titles and abstracts in MEDLINE (including recent un-indexed papers listed in Pre-MEDLINE), EMBASE and the Cochrane Library with no language or date restrictions.

The search of Medline searched for type 2 diabetes in the thesaurus and also in free text. Relevant statistical terms in the thesaurus search were combined with the word *risk* adjacent (within three words) of other key words in the search. The relevant statistical terms and the risk search were combined with the type 2 diabetes search. In order to narrow the search further the function to focus on *diabetes* or *prediabetic state* as the main focus of the article was chosen. This result was incorporated to produce the final result. This resulted in 6,169 papers. The search of EMBASE searched for type 2

diabetes in the thesaurus and also in free text. Relevant statistical terms in the thesaurus search were combined with the word *risk* adjacent (within three words) of other key words in the search. The relevant statistical terms and the risk search were combined with the type two diabetes search. In order to narrow the search further the function to focus on *non insulin dependent diabetes mellitus* as the main focus of the article was chosen. This result was incorporated to produce the final result. This resulted in 6,947 papers. The text word strategy was used only for unindexed MEDLINE papers resulting in 524 papers. A MESH search only was performed in the Cochrane Library for type 2 diabetes mellitus and risk resulting in 716 titles. Details of the exact search strategies executed can be found in Appendix 3.

From the search strategy 14,356 titles were moved into the electronic reference package ENDNOTE and duplicates automatically removed. This resulted in a total of 10,275 titles. Some duplicates remained which ENDNOTE was not able to discriminate between and these were removed manually resulting in 8,864 titles.

Two independent researchers (DN and Rohini Mathur (RM)) independently scanned all 8,864 titles and if it was suspected that the title represented a paper which met the inclusion and exclusion criteria, the abstracts were reviewed.

**Initial inclusion criteria**

**Study design:** Any study deriving or validating a risk model or score for type 2 diabetes.

**Population:** Adults over age 18 with no upper age limit.

**Intervention:** Developing models or scoring systems based on type 2 diabetes risk factors to predict temporal risk of type 2 diabetes in adults and/or validation of a type 2 diabetes risk model or score.

**Outcomes:** Any relevant predictive outcomes, including discrimination and calibration.


**Initial exclusion criteria**

1. Studies which had not finished recruiting.

2. Studies examining one or more single risk factors that had not been linked together to form a model or scoring system.

3. Screening and early detection studies.

4. Models predicting genetic mutations rather than type 2 diabetes.

5. Case series.

6. Studies carried out on animals.


Title scanning was finished in March 2011. DN marked 141 titles as potentially meeting inclusion criteria and RM marked 124 titles. Where possible, abstracts were reviewed, if not, full papers were requested. Following available abstract review DN requested 79 full papers for further review and RM 55 full papers. These were combined and 38 duplicates removed resulting in 96 full papers. Owing to the methodologically complex nature of this dataset DN categorised these 96 papers within

an Excel spreadsheet under the following headings, as it was becoming apparent that the initial inclusion and exclusion criteria were not fully adequate and that there were major methodological problems in the literature:

1. First author

2. Journal

3. Year of publication

4. Country of origin of research population

5. Language

6. Brief context

7. Research performed on adults without preselected diseases or risks– yes/no

8. Derived a risk score – yes/no

9. Validated the risk score on a separate population

10. Type of study – e.g. cross-sectional or cohort

11. Period of time derivation population studied (years)

12. Age range of derivation group (years)

13. Period of time validation population studied (years)

14. Age range of validation group (years)

15. Method of confirming diagnosis in derivation group

16. Method of confirming diagnosis in validation group

17. Calibration statistics presented

18. Discrimination statistics used

19. Validated or compared or used other scores

On reviewing this spreadsheet, and after discussion with the project team, the inclusion and exclusion criteria were modified as below. This reduced the number of final papers included in the review, but would not have influenced selection of titles, abstracts or full papers prior to this stage, as the initial inclusion and exclusion criteria was broader than the final criteria. The main difference was the exclusion of cross-sectional studies as these were considered to be less methodologically robust.

**Final inclusion criteria**

**Study design:** Any cohort study which derives and/or validates a type 2 diabetes risk model or score.

**Population:** Adults over age 18 with no upper age limit.

**Intervention:** Developing models or scoring systems for type 2 diabetes based on regression analysis to predict temporal risk of type 2 diabetes in adults and/or validation of a type 2 diabetes risk model or score, on a different population.

**Outcomes:** Any relevant predictive outcomes, including discrimination and calibration.

**Final exclusion criteria**

**Study design:** Cross-sectional designs, other screening or early detection studies, genetic studies or case series.

**Population:** Pre-selected populations with existing risk factors or disease, studies on under 18s, studies carried out on animals, and studies which have not finished recruiting.

**Intervention:** Studies examining one or more single risk factors that have not been linked together to form a model or scoring system, and studies that applied a known risk model or score to a population but did not evaluate its statistical potential.

Ten papers out of the 96 full papers were in a non-English language (Persian, Chinese, Dutch, German (4), Polish, Hungarian and Spanish). Although three had no English abstract it was obvious to both DN and RM that they were either a commentary, editorial or article, and did not contain quantitative data or a risk model or score. These were excluded. Four further papers were excluded after discussion because the English abstract clearly indicated they did not meet the inclusion criteria. Three were marked for translation.

Following independent full paper review, using the final inclusion and exclusion criteria, DN reduced the total to 24 full papers and RM to 22. One paper selected for inclusion by RM was excluded after joint discussion; her other 21 choices were within DNs final set. The extra three papers selected for inclusion by DN were included after discussion and in particular, revisiting the inclusion criteria. This resulted in 24 full papers. Therefore 69 out of 96 papers were excluded after full paper review from the main literature search, leaving 27 (including 3 marked for translation) remaining.

**Full papers from other sources**

Seven more papers were added from the initial scoping search, one paper from a Google search, and eight papers following a Google citation review of the majority of the 96 full papers (papers that did not have citation tracking included foreign language papers and those that were of limited relevance to diabetes risk models or scores). DN performed these searches alone. This resulted in 16 further full papers.

DN and RM divided the task of reference saturation from key references, which included all 43 papers included for data extraction to this point, other papers in the 96 full papers and also commentaries, editorials, PhD theses and other relevant sources. This added three full papers.

In total 46 papers were selected for full data extraction.

**Quantitative data extraction**

Quantitative data extraction was undertaken under the following headings:

1. Author

2. Journal

3. Year

4. Country/Context

5. Language

6. How many models?

7. Name of study derived from e.g. Framingham

8. Name of score e.g. QDScore

9. Study design

10. If validation only study where is it derived from?

11. Reasons for original recruitment

12. Sample size

13. Study duration

14. Year studied from - to

15. Age

16. Gender

17. Ethnicity

18. Deprivation

19. Qualitative descriptor of cohort / sample

20. Diabetes excluded in inception cohort - yes/no

21. How diabetes diagnosed at inception

22. How diabetes diagnosed at intervals +/or completion

23. Follow-up rate to end of cohort

24. Prevalence of diabetes at cohort inception

25. Incidence of diabetes at end of cohort

26. Risk score components

27. Sensitivity (of authors preferred cut-off - note rationale)

28. Specificity (of authors preferred cut-off - note rationale)

29. How derived cut-off score (s)

30. AUROC + 95% CI

31. Positive predictive value (PPV)

32. Negative predictive value (NPV)

33. Calibration

34. Other statistics

35. Description of internal validation if present

36. Percentage that would need further testing if classified high risk by authors

37. Extra notes

Some studies offered numerous different models with different risk factors included in each. It was beyond our capacity to study in detail every one of these models, and often

the authors themselves had concluded that, for example, of six models tested one clearly outperformed the others. Where studies had tested multiple models or scores with minimal difference in the risk factors included, we extracted data from the authors' preferred model(s) or (if no preferences were stated in the paper) the ones judged by two researchers to be the most complete or statistically robust. Our aim was to use statistical meta-analysis where appropriate and to present heterogeneous data in disaggregated form.

DN did primary data extraction for 23 full papers; CM did 10 and RM did 9. RM and DN discussed one of the included papers which transpired not to have been peer reviewed. This was then excluded, reducing the total of included papers to 45. DN worked with translators to extract data from the three papers in Persian, Spanish and Dutch.

Dr Tom Dent (TD) joined the project team to independently double check all of the quantitative data extraction apart from the translated papers. As the translations were done by DN working alongside a GP, medical student or scientist, these papers had already been double checked. TD raised concerns that two further papers did not meet the criteria and this was discussed with the entire research team and were excluded. The final number of papers included was therefore 43.

**Qualitative data extraction**

Along with Professor Trisha Greenhalgh (TG), and drawing on the principles of realist review (a relatively new form of systematic literature review which uses mainly

qualitative methods to produce insights about the interaction between context, mechanism and outcome - explaining instances of both success and failure[49]) the following research question was addressed:

*What is the relationship between the components of the score, the context in which it was intended to be used and the mechanism by which it might improve outcomes for patients?*

Data was extracted and entered on a spreadsheet under seven headings:

1. *Intended users:* Authors' assumptions (if any) about who would use the risk score, on which subgroups or populations

2. *Proposed action based on the score result:* Authors' assumptions (if any) on what would be offered to people who score above the designated cut-off for high risk

3. *Mechanism:* Authors' hypothesised (or implied) mechanism by which use of the score might improve outcomes for patients

4. *Descriptor:* Authors' adjectives to describe their risk model or score

5. *Relative advantage:* Authors' claims for how and in what circumstances their model or score outperforms previous ones

6. *Concerns:* Authors' stated concerns about their model or score

7. *Real world use, including citation tracking:* Actual data in this paper or papers citing it on use of the score in the real world

TG performed the data extraction. DN played a supporting role in this part of the systematic review (as oppose to leading the quantitative part of the review). For the realist review, DN independently double checked the data extraction from 15 of the 43 papers, agreed with the findings and added a small amount of additional material. Context-mechanism-outcome interactions hypothesised or implied by authors were discussed among our research team and these mechanisms were explored further by re-reading the full sample of papers with all emerging mechanisms in mind.

**Impact analysis**

We assessed the impact of each risk score in our final sample using three criteria: [a] any description in the paper of use of the risk score beyond the population in which it was developed and validated, [b] number of citations of the paper in Google Scholar, and number of these which described use of the risk score in an impact study (TG did a second full citation track of the included papers looking for this specifically), and [c] critical appraisal of any impact studies identified on this citation track. In this phase, we were guided by the question:

*What is the evidence that this risk score has been used in an intervention which improved (or sought to improve) outcomes for individuals at high risk of diabetes?*

**Prioritising papers for reporting in the published paper**

Given the large number of papers and risk models and scores in our final sample, we decided for clarity to highlight the seven models or scores most likely to be useful to

practising clinicians, public health specialists or lay people. Adapting Altman et al's quality criteria for risk scores,[37] the following were used to guide the prioritisation of scores for reproduction in a concise and easily accessible table: [a] external validation by a separate research team on a different population (*generalisability*), [b] statistically significant *calibration,* [c] a *discrimination* greater than 0.70, and [d] 10 or fewer components (*usability*). External validations frequently altered the original score by either: [a] the number of risk factors, [b] categorisation of the risk factor e.g. using a different ethnicity, and [c] not stating exactly which risk factors they used from the original model or score. We did not exclude the external validation on this basis as part of our prioritisation was linked to impact. One score which has not yet been externally validated was included in the concise table because the review of impact in the qualitative part of the review had highlighted it.

**Chapter 4: Results and Discussion**

Below I describe the main findings of the systematic review. The full version of this systematic review, already published in the BMJ, can be found in this thesis as Appendix 4. Only one table has been reproduced in the main body of the thesis. All other tables can be found in Appendix 4 and are referred to in the text in this section for reference purposes only.

**Results**

Figure 2 shows the flow of papers through the study.



Records identified through search of databases (n=14 356)

Records after duplicates removed (n=8864)

Records screened (n=8864)

Records excluded (n=8768)

Full text records (n=96)

Additional records identified through other sources (n=19)

Full text articles assessed for eligibility (n=115)

Full text articles excluded (n=72):
 Cross sectional design (n=26)
 Study population had preselected risk factors or other disease (n=7)
 Failed to derive or validate risk model or score (n=29)
 Children and adolescents studied (n=1)
 Genetic study (n=1)
 Risk score not validated in terms of discriminatory value, simply ran
  a risk score on a cross sectional population (n=3)
 Derived score according to literature, not by regression analysis (n=1)
 Not a research paper (n=4)

Studies included in review (n=43)

Models or scores trialed (n=145)
Models or scores selected (n=94)

Models or scores with potential for use in clinical practice (n=7)

**Figure 2: Flow of studies through the systematic review**

In summary, 115 papers were analysed in detail to produce a final sample of 43 - 18 of which described development of one or more risk models and/or scores;[32] [50-66] 17 described external validation of one or more risk models and/or scores on a new population(s);[67-83] and eight did both.[33] [39] [84-89] In all, the 43 papers described 145 models, from which we selected 94 for full data extraction (the other 51 were minimally different, were not the authors' preferred model, or lacked detail or statistical robustness). This sample of 94 included 55 derivations of risk-predicting models on a base population and 39 external validations of 14 different risk models or scores on new populations. Studies were published between 1993 and 2011, but most have appeared in the past three years. [Figure 3] Indeed, even given that weaker cross-sectional designs had been excluded, findings suggested that new diabetes risk models and scores were currently being published at a rate of approximately one every three weeks.
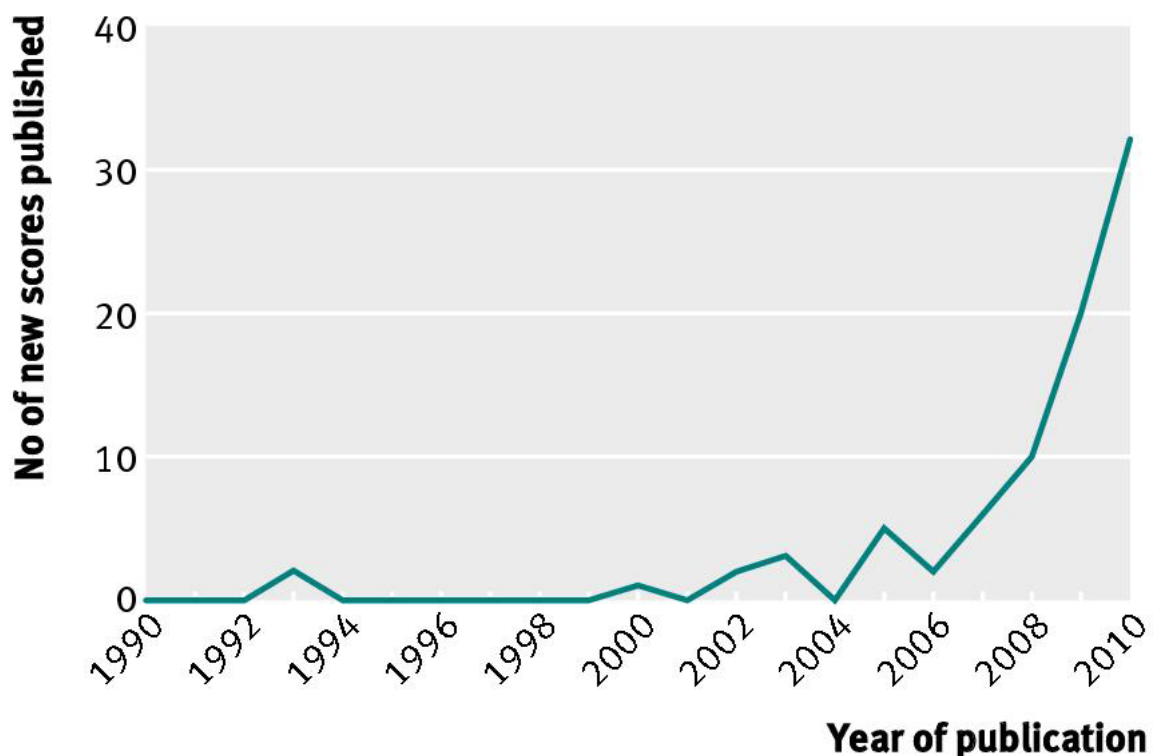


**Figure 3: Publication of diabetes risk models and scores 1990-2010.**
*Eleven new risk models and scores had been published in the first five months of 2011*

Appendix 4 (Table 1) gives full details of the studies in the sample, including the origin of the study, setting, population, methodological approach, duration, and how diabetes was diagnosed. In sum, the studies were highly heterogeneous. Models were developed and validated in 17 countries representing six continents (30 in Europe, 25 in North America, 21 in Asia, 8 in Australasia, 8 in the Middle East, one in South America and one in Africa).

Comparisons across studies were problematic owing to heterogeneity of data and highly variable methodology, presentation techniques and missing data. Cohorts ranged in size from 399 to 2.54 million. The same datasets were often used in several different models in the same paper. Ten research populations were used more than once in different papers.[32 51 56 57 59 61-63 65-71 78 80-84 87 88] In total, risk models were tested on 6.88 million participants, although this figure includes multiple tests on the same datasets. Participants aged 18 to 98 years were studied for periods ranging from 3.15 to 28 years. Completeness of follow-up ranged from 54 to 99% and incidence of diabetes across the time periods studied ranged from 1.3 to 20.9%.

None of the risk scores in the sample was developed on a cohort recruited prospectively for the purpose of devising a model or score. Rather, all authors used the more pragmatic approach of retrospectively studying a research dataset which had been assembled some years previously for a different purpose. There were 42 studies that excluded known diabetes in the inception cohort. Diagnosis of diabetes at inception and completion of cohort was done in different ways including self-report, patient questionnaires, clinician diagnosis, electronic code, ICD code, disease or drug registers, diabetes medication, dietary treatment, fasting glucose, oral glucose tolerance test,

Glycated Haemoglobin A1c (HbA1c), and some studies did not state the method. Half the studies used different diagnostic tests at inception and completion.

One-third of the papers focused almost exclusively on the statistical properties of the models or score(s) reported. Many of the remainder had a clinician (diabetologist or general practitioner) as co-author and included an (often short and speculative) discussion on how the findings might be applied in clinical practice. Three described their score as a 'clinical prediction rule'.[60 66 74]

Details of the components of the 94 risk scores included in our final sample and their statistical properties − including (where reported) their discrimination, calibration, sensitivity, specificity, positive and negative predictive value, and AUROC − are shown in Appendix 4 (Table 2). Many papers offered additional sophisticated statistical analysis, though there was no consistency between research teams in the approach or statistical tests used. Heterogeneity of data (especially demographic and ethnic diversity of validation cohorts and different score components) in the primary studies precluded formal meta-analysis.

All 94 risk scores presented a combination of risk factors as significant in the final model, and different models weighted different components differently. The number of components in a single risk score varied from 3 to 14 (n=84, mean 7.8, SD 2.6). The seven scores which we classified as having high potential for use in practice offered broadly similar components and had similar discriminatory properties (AUROC 0.74-0.85). These seven highlighted scores are show in Table 3.

| SCORE/STUDY, COUNTRY, AUTHOR/YEAR | RISK FACTORS INCLUDED IN SCORE | AUROC | CALIBRATION | EXTERNAL VALIDATION | | |
|---|---|---|---|---|---|---|
| | | | | Year, Country | AROC | CALIBRATION |
| ARIC, Germany (Atherosclerosis Risk in Communities; Schmidt 2005)[61] | Age, ethnicity, waist circumference, height, systolic blood pressure, family history of diabetes, fasting plasma glucose levels, triglyceride levels, high density lipoprotein cholesterol levels | 0.80 | Not stated | 2010,[76] USA | 0.84 | HL p<0.001, after recalibration p>0.10 |
| Ausdrisk, Australia (Chen 2010)[51] | Age, sex, ethnicity, parental history of diabetes, history of high blood glucose, use of antihypertensive drugs, smoking, physical inactivity, waist circumference | 0.78 | HL p=0.85 | Not externally validated but has been studied as part of an intervention to improve outcomes[90] | | |
| Cambridge Risk Score, UK (Rahman 2008)[80] | Age, sex, use of current corticosteroids, use of antihypertensive drugs, family history of diabetes, body mass index, smoking | 0.74 with threshold of 0.38 | Not stated | 2010,[81] UK* | 0.72 | HL P=0.77 |
| FINDRISC, Finland (Lindstrom 2003)[33] | Age, body mass index, waist circumference, use of antihypertensive drugs, history of high blood glucose, physical inactivity, daily consumption of vegetables, fruits, and berries | 0.85 | Not stated | 2010,[68] Holland, Denmark, Sweden, UK, Australia* | 0.76 | HL p =0.27 |

| Framingham Offspring Study, USA (Wilson 2007)[66] | Fasting plasma glucose levels, body mass index, high density lipoprotein cholesterol levels, parental history of diabetes, triglyceride levels, blood pressure | 0.85 | Not stated | 2010,[76] USA | 0.78 | HL p<0.001, after recalibration p>0.10 |
|---|---|---|---|---|---|---|
| San Antonio Risk Score, clinical model, USA, (Stern 2002)[32] | Age, sex, ethnicity, fasting plasma glucose levels, systolic blood pressure, high density lipoprotein cholesterol levels, body mass index, family history of diabetes in first degree relative | 0.84 | HL P>0.2 | 2010,[76] USA | 0.83 | HL p<0.001, after recalibration p>0.10 |
| | | | | 2010,[70] Iran * | 0.83 | HL p≤0.001 after recalibration P=0.131 |
| | | | | 2010,[81] UK* | 0.78 | HL P=0.42 |
| | | | | 2010,[84] Iran* | 0.78 | HL P=0.264 |
| QDScore, UK (Hippisley-Cox 2009)[54] | Age, sex, ethnicity, body mass index, smoking, family history of diabetes, Townsend deprivation score, treated hypertension, cardiovascular disease, current use of corticosteroids | 0.83 men, 0.85 women | Brier score 0.078 men, 0.058 women | 2011,[72] UK | 0.80 men, 0.81 women | Brier Score 0.053 men, 0.041 women |

**Table 3: Components of seven diabetes risk models or scores with potential for adaptation for use in routine clinical practice**
*Validation used more, less or substituted risk factors from the original risk score or did not state the exact factors it used.*

These seven validated diabetes risk scores were judged to be the most promising for use in clinical or public health practice. The judgments on which this selection was based were pragmatic; other scores not listed in Table 3 could possibly prove more fit for purpose in certain situations and settings. One score that had not yet been externally validated according to the pre-set criteria was included in Table 3 as it was already being incentivised in a national diabetes prevention policy.[91] This Australian government scheme is targeted at middle aged adults 40-49 years (15-54 years for Aboriginal or Torres Strait Islander peoples), and attracts a Medicare rebate. High risk individuals undergo further assessment by their GP which may involve a reduced cost lifestyle intervention. This risk score was also included in the qualitative part of the review for studies using diabetes risk models or scores as part of an intervention to improve outcomes (Appendix 4, Table 5).[90] Subsequent correspondence with the author revealed a small validation study on approximately 500 women who were part of a separate osteoporosis study, that had been reported in a letter to the Medical Journal of Australia.[92]

Overall, AUROCs ranged from 0.60 to 0.91. Certain components used in some scores (e.g. biomarkers) are rarely available in some pathology laboratories and potentially too expensive for routine use. Some models which exhibited good calibration and discrimination on the internal validation cohort performed much less well when tested on an external cohort[79 85] suggesting that the initial model may have been overfitted by inclusion of too many variables that had only minor contributions to the total risk.[93] Overfitting arises because any given risk factor in the score does not occur very frequently

in the original derivation study. This manifests itself most noticeably when the score is used on a different population, and produces incongruent results.[94] Whilst genetic risk scores were not sought out, those studies which had included genetic markers alongside socio-demographic and clinical data all found that the former added little or nothing to the overall model.[50 64 78 81]

Reporting of statistical data in some studies was very incomplete. For example, only 40 of the 94 models quantified any form of calibration statistic. There were 43 which presented sensitivity and specificity, 27 justified the rationale for cut-off points, 22 presented a positive and 19 a negative predictive value, and 26 made some attempt to indicate the percentage of the population that would need clinical follow-up or testing if they scored 'high risk'. Some models performed poorly (e.g. there was a substantial gap between expected and observed numbers of participants who developed diabetes over the follow-up period). The false positive and false negative rates in many scores raised serious questions about their utility in clinical practice (e.g. positive predictive value ranged from 5 to 42%; negative predictive value ranged from 88 to 99%). However, some scores were designed as non-invasive preliminary instruments with a recommended second phase involving a blood test.[39 58 67 68 70 73 83] Risk models and scores tended to 'morph' when they were externally validated because research teams dropped components from the original (for example, if data on these were not available), added additional components (for example, to compensate for missing categories), or modified what counted in a particular category (for example, changing how ethnicity was classified); in some cases these modifications were

not clarified.  Since a key dimension of implementation is appropriate adaptation to a new context, this probably did not negate the external validation.

The qualitative findings from the risk scores are shown in Appendix 4 (Table 3).  Of the 43 papers in the full sample, three did not recommend use of the model tested because the authors felt it had no advantage over existing approaches.[64 71 75]  Authors of the other 40 papers considered that at least one of their score(s) should be adopted and used, and made various claims to justify this.  The commonest adjective used by authors to describe their score was 'simple' (26 out of 43); others included 'low-cost', 'easily implemented', 'feasible' and 'convenient'.

In total, 16 of the 43 studies which recommended use of a particular risk model or score did not designate an intended user for it.  Some authors assigned agency to a risk score (i.e. they stated, perhaps inadvertently, that *the score itself* had the potential to prevent diabetes, change behaviour or reduce health inequalities).  Whilst most authors did state an intended target group, this was usually given in vague terms (e.g. 'the general population' or 'individuals who are likely to develop diabetes in the near future').  Eleven of the 43 papers gave a clear statement of what intervention might be offered, and by whom, to people who scored above the cut-off for high risk.  The other papers made no comment on this or used vague terms such as 'preventive measures' without specifying who would deliver these.

In all, the authors of the papers in our full sample either explicitly identified or appeared to assume ten mechanisms by which (singly or in combination) use of the diabetes risk scores might lead to improved patient outcomes:

1. **Clinical**

Direct impact - clinicians will pick up high risk patients during consultations and offer advice that leads to change in patients' behaviour and lifestyle.

Indirect impact - routine use of the score increases clinicians' awareness of risk for diabetes and motivation to manage it.

2. **Self assessment**

Direct impact - people are alerted by assessing their own risk (for example, using an online tool), directly leading to change in lifestyle.

Indirect impact - people, having assessed their own risk, are prompted to consult a clinician to seek further tests or advice on prevention.

3. **Technological**

Individual impact - a risk model programmed into the electronic patient record generates a point of care prompt in the clinical encounter.

Population impact - a risk model programmed into the electronic patient record generates aggregated data on risk groups, which will inform a public health intervention.

4. **Public health**

Planners and commissioners use patterns of risk to direct resources into preventive healthcare for certain subgroups.

### 5. Administrative

An administrator or healthcare assistant collects data on risk and enters these onto the patients' records, which subsequently triggers the technological, clinical, or public health mechanisms.

### 6. Research into practice

Use of the risk score leads to improved understanding of risk for diabetes or its management by academics, leading indirectly to changes in clinical practice and hence to benefits for patients.

### 7. Future research

Use of the risk score identifies focused subpopulations for further research (with the possibility of benefit to patients in later years).

Risk models and scores had been developed in a wide range of different health systems. Differences in the components of the scores could be explained partly in terms of their intended context of use (which in turn were specific to the setting and health system).

None of the 43 papers that validated one or more risk scores described the actual use of that score in an intervention phase. Furthermore, whilst these papers had been cited by a total of 1,883 (range 0-343, median 12) subsequent papers, only nine of those 1,883 papers (listed in Appendix 4 (Table5)) described application and use of the risk score as part of an impact study aimed at changing patient outcomes. These covered seven studies, of which three had reported definitive results. All three reported positive changes in individual risk

factors, but surprisingly none recalculated participants' risk scores after the intervention period to see if they had changed. Whilst one report on the ongoing FIN-D2D study suggests that incident diabetes has been reduced in 'real world' (i.e. non-trial) participants who were picked up using a diabetes risk score and offered a package of preventive care,[95] this is a preliminary and indirect finding based on drug reimbursement claims, and no actual data are given in the paper. With that exception, no published impact study on a diabetes risk score has yet demonstrated a reduction in incident diabetes.

**Discussion**

The lengthy background to this section of the thesis in chapter 4 highlighted the complexity of type 2 diabetes risk models and scores. A brief review of the literature by a public health specialist or general practitioner is unlikely to result in quick translation of risk scores into everyday practice. A systematic review was needed to make sense of this field of research. Methodology for the review as described in Chapter 4 was complex and evolved iteratively to take account of the diversity of diabetes risk scores.

The results of the systematic review have demonstrated that a small number of diabetes risk models and scores exist based on data that are readily available and which provide a good but not perfect estimate of the chance that a non-diabetic adult will develop diabetes in the medium-term future. A few research teams have undertaken exemplary development and validation of a robust model, reported its statistical properties thoroughly, and followed through with studies of impact in the real world.

Included studies were not entirely free from bias and confounding. This is because the 'pragmatic' use of a previously assembled database or cohort to develop or validate a diabetes risk score brings an inherent selection bias, as described previously. For example, the British Regional Heart Study cohort was selected to meet the inclusion criteria for age and co-morbidity defined by its original research team and oriented to research questions

around cardiovascular disease; the population for the QDScore is drawn from general practice records and hence excludes those not registered with a GP).

Most papers in the final sample of 43 papers had one or more additional limitations. They: [a] reported risk models or scores that required collection of data not routinely available in the relevant health system, [b] omitted key statistical properties such as calibration and positive and negative predictive values that would allow a clinician or public health commissioner to judge the practical value of the score, or [c] omitted to consider who would use the score, on whom and in what circumstances. It was identified that there was a mismatch between the common assumption of authors who develop a risk model (that their 'simple' model can now be taken up and used) and the actual uptake and use of such models (which seems to happen very rarely). However, there has recently been an encouraging − if limited − shift in emphasis from the exclusive pursuit of statistical elegance (e.g. maximising AUROC) to undertaking applied research on the practicalities and outcomes of using diabetes risk scores in real-world prevention programmes.

The strengths of the systematic review are: [a] use of mixed methodology, [b] orientation to patient-relevant outcomes, [c] extraction and double-checking of data by five researchers, and [d] inclusion of a citation-track to identify recently published studies and studies of impact. Both standard systematic review methods (to undertake a systematic and comprehensive search, translate all non-English texts, and extract and analyse quantitative data) and realist methods (to consider the relationship between the components of the score,

the context in which it was intended to be used and the mechanism by which it might improve outcomes for patients) were employed.

The main limitation of the review is that data techniques and presentation in the primary studies varied so much that it was problematic to determine reasonable numerators and denominators for many of the calculations. This required pragmatic decisions to be made to collate and present data as fairly and robustly as possible while also seeking to make sense of the vast array of available scores to the general medical reader. It is recognised that the final judgement on which scores are, in reality, easy to use will lie with the end-user in any particular setting. Secondly, authors of some of the primary studies included in this review were developing a local tool for local use and made few or no claims that their score should be generalised elsewhere. Yet, the pioneers of early well-known models[32][33] have occasionally found their score being applied to other populations (perhaps ethnically and demographically different from the original cohorts), their selection of risk factors being altered to fit the available categories in other datasets, and their models being re-calibrated to provide better goodness of fit. All this revision and recalibration to produce 'new' scores makes the systematic review of such scores at best an inexact science.

After finalising the systematic review, and prior to publication in the BMJ, two separate systematic reviews were published by teams in Cambridge[96] and Oxford.[94] These reviews applied a different but complementary approach.

The Cambridge Review was undertaken by long term experts in diabetes risk scores, including the original author of the Cambridge Diabetes Risk Score, which is world renowned.[31] Although this score was created from cross-sectional data, it was later used in a cohort study by Rahman et al in 2008[80] and was hence included in the systematic review included within this thesis (the London Review). The Cambridge Review's aim was threefold: [a] to identify scores for diabetes risk prediction, [b] to evaluate performance in a new population, of scores derived and validated elsewhere, and [c] analysing methodological difficulties. They also only included cohort studies of adults that excluded persons with diabetes at cohort inception. Like my colleagues and I, they also had to make pragmatic decisions in a complex field. For example, they decided where scores appeared in more than one paper, to include the paper that had the most detail on that score's performance, as oppose to the approach in the London Review which resulted in analysis of 39 external validations of 14 models or scores. Although they collected similar statistical information on each score (including calibration) it was of note that in the main table in the published paper they chose only to present discrimination. Whilst this is understandable as it is the statistic that is most often complete, it is only part of the picture of performance as described in Chapter 4. Comparison of discrimination between scores as new risk factors are added and subtracted, and different population groups experimented on, has clearly been a large part of this research field and the academic competition between different teams. As experienced researchers the Cambridge Review in this sense represents the expected review from the leaders of this field. The results catalogued a list of descriptive statistics without meta-analysis. A qualitative assessment of real world impact was not

performed (and not expected given the background of the research team). The Cambridge team's conclusions included: [a] that risk score performance differed in new populations when derived from studies with varying ethnic structures, [b] that there were scores that used easily obtainable data, [c] that identifying high risk groups could help with targeted prevention efforts, and [d] that scores which included glycaemic indices (or other blood tests) performed better.

By contrast, the Oxford Review was carried out by an experienced research team without specific expertise in diabetes risk scores. This same team also externally validated the QDScore on a new population[72], and are developing a track record of validating other risk scores derived from the creators of the QDScore.[97-100] The Oxford Review had a different focus to both Cambridge and London approaches, reviewing primarily methods rather than performance. They also reviewed scores for undiagnosed diabetes, hence their review is not strictly comparable to the London and Cambridge Reviews. Appropriately, they extracted fields from their dataset including coding and model-building strategies. And like the other two reviews they performed no meta-analysis. In their findings section, these authors focused on the poor and unstandardised methodology of models and scores.

Together the three reviews cover methodological, statistical, and 'application in the real world weaknesses', and whilst there is some duplication, the different angles of the reviews make them complementary to addressing different aspects of risk models and scores. All three reviews highlight the complex and disparate nature of risk models and scores, and

pave the way for a standard approach to derivation, validation and application, which is, at present, lacking.

The finding that diabetes risk scores appear to be rarely used can be considered in the light of the theoretical literature on diffusion of innovation. As well as being a mathematical model, a risk score can be thought of as a complex, technology-based innovation, the incorporation of which into business as usual (or not) is influenced by multiple contextual factors. This includes the attributes of the score in the eyes of potential adopters (relative advantage, simplicity and ease of use); adopters' concerns (including implications for personal workload and how to manage a positive score); their skills (ability to use and interpret the technology); communication and influence (e.g. whether key opinion leaders endorse it); system antecedents (including a healthcare organisation's capacity to embrace new technologies, workflows and ways of working); and external influences (including policy drivers, incentive structures and competing priorities).[101 102]

Whilst the *developers* of most diabetes risk scores are in little doubt about their score's positive attributes, this confidence seems not to be shared by practitioners, who may doubt the accuracy of the score and/or the efficacy of risk modification strategies. Measuring diabetes risk competes for practitioners' attention with a host of other tasks, some of which bring financial and other rewards. Furthermore the very low positive predictive values may spell trouble for commissioners. Identifying a person as '[possibly] high risk' will inevitably entail a significant cost in clinical review, blood tests and (possibly) intervention

and follow-up. Pending the results of ongoing impact studies, this may not be the best use of scarce resources.

Whilst most authors of papers describing diabetes risk scores have hypothesised (or appear to have assumed) a *clinical* mechanism of action (i.e. that the score would be used by the individual's clinician to target individual assessment and advice), the limited data available on impact studies suggest that a particularly promising area for further research is interventions which prompt self-assessment (i.e. lay people measuring their own diabetes risk). Risk scores which rely entirely on such questions may be hosted on the Internet (see for example http://www.diabetes.org.uk/riskscore). Some researchers have used self-completion postal questionnaires as the first part of a step-wise detection programme.[103] To the extent that these instruments are valid, they can identify two types of individual: [a] those who already have diabetes whether they know it or not (hence, the questionnaire may serve as a self-administered screening tool for undiagnosed diabetes) and [b] those at high risk of developing diabetes (hence, it may also serve as a prediction tool for future diabetes). Hence, diabetes prevalence rates derived from self-assessment studies cannot be compared directly with the incident diabetes rate in a prospective longitudinal sample from which those testing positive at baseline have been excluded.

The findings did not support the recommendation of a single, preferred diabetes risk score. There is no universal 'ideal' risk score, since the utility of any score depends not merely on its statistical properties but also on its context of use, which will also determine which

78

types of data are available to be included.[41 104] Even when a risk model has excellent discrimination (and especially when it does not) the trade-off between sensitivity and specificity plays out differently depending on context.

Listed below are some suggested questions a public health specialist, general practitioner or commissioner could ask when faced with the dilemma of which risk score to use:

**What is the intended use case for the score?**

If intended for use:

> In clinical consultations - the score should be based on data on the medical record
>
> For self assessment by lay people - the score should be based on things a layperson would know or be able to measure.
>
> In prevention planning - the score should be based on public health data.

**What is the target population?**

If intended for use in high ethnic and social diversity, a score that includes these variables may be more discriminatory.

**What is expected of the user of the score?**

If for opportunistic use in clinical encounters, the score must align with the structure and timeframe of such encounters and competencies of the clinician, and (ideally) be linked to

an appropriate point of care prompt. Work expected from the intended user of the score may need to be incentivised or remunerated, or both.

**What is expected of the participants?**

If to be completed by lay people, the score must reflect the functional health literacy of the target population.

**What are the consequences of false positive and false negative classifications?**

In self completion scores, low sensitivity may falsely reassure large numbers of people at risk and deter them from seeking further advice.

**What is the completeness and accuracy of the data from which the score will be derived?**

A score based on automated analysis of electronic patient records may include multiple components but must be composed entirely of data that are routinely and reliably entered on the record in coded form, and readily searchable (thus, such scores are only likely to be useful in areas where data quality in general practice records is high).

**What resource implications are there?**

If the budget for implementing the score and analysing data is fixed, the cost of use must fall within this budget.

**Given the above, what would be the ideal statistical and other properties of the score in this context of use?**

What trade-offs should be made (sensitivity *v* specificity, brevity *v* comprehensiveness, one stage *v* two stage process)?

Millions of participants across the world have already participated in epidemiological studies aimed at developing a diabetes risk score. There is now a menu of possible scores available to those who seek to use them and/or validate them in new populations, none of which is perfect but all of which have strengths.

My initial impression of diabetes risk scores being too numerous and too inaccessible for GPs or public health specialists to readily decide which one to use in practice, was to an extent allayed by the findings from the systematic review. It appeared that in a vast and methodologically complex area, suitable risk scores which had potential to be readily adapted for any given clinical or public health uses were available. I was satisfied that performance had been documented comprehensively, and most importantly impact assessed, revealing that this was mostly absent.

The next challenge was to take a readily accessible and usable score and study the feasibility of using it in real life practice. With this conclusion from the systematic review, I moved to the next stage of this research and chose one of the scores to use on the local population in East London, presented in Section 3.

**SECTION 3: Cross-Sectional Analysis of Electronic General Practice Records**

**Chapter 5: Background, Aims, Methods**

In this section, in the context of the epidemic of type 2 diabetes described in Section 1, and the availability of some suitable risk scores for use in clinical and public health practice from the review in Section 2, I sought to translate and apply this research in a real world setting. Having lived and worked in East London for some years I was already aware of the great needs for prevention of diabetes and its common risk factors, such as obesity. In this section I will describe the setting and population of East London, how I selected one of the risk scores to use on local electronic general practice records, outline the aims and methods of an empirical cross-sectional study, and present selected results with accompanying discussion. This was published as a special report on diabetes risk for local NHS partners and is available on the Queen Mary University London website. The academic paper has been submitted to the British Journal of General Practice and is currently being peer reviewed. This can be found in Appendix 5.

**Background**

Inner North East London comprises the boroughs of Tower Hamlets (population 241,747)[26], Newham (population 265,688)[23] and City (population 9,502) & Hackney (population 229,036).[24] These make up three PCTs as City & Hackney are grouped together. The estimated population size overall using Greater London Authority estimates

from the most recent Joint Strategic Needs Assessments and the Tower Hamlets Population

Change and Growth Model is 745,973.

East London is ethnically diverse and of low socioeconomic status compared to England as a whole. Joint Strategic Needs Assessments reveal non-white ethnic groups make up approximately 50% of the population in Tower Hamlets, 40% in Hackney, 20% in the City and 70% in Newham. Certain ethnic groups suffer low-literacy and obesity. The combination of ethnicity (generational and genetic risk) and cultural/linguistic barriers combined with poverty, increase the risk of diabetes significantly. The three areas ranked 3rd (City&Hackney), 4th (Tower Hamlets) and 7th (Newham) for highest Index of Multiple Deprivation (IMD) in England in 2007.[105] Tower Hamlets moved down to 8th place in the 2010 figures as shown in Table 4.[106]

| Primary Care Trust | Average IMD Score 2010 (higher is more deprived) | National rank of average score (higher ranking is more deprived – out of a total of 152 PCTs) |
|---|---|---|
| Newham | 41.84 | 3rd |
| City and Hackney | 41.28 | 4th |
| Tower Hamlets | 39.59 | 8th |

*For comparison average Index of Multiple Deprivation score in 2010 in England was 21.67.*
**Table 4: East London Index of Multiple Deprivation 2010**

The registered general practice population is significantly higher than the geographical population due to people remaining on lists after leaving the area and people who are registered but do not live in the boroughs. When the cross-sectional study took place the

general practice registered population was 881,896 (135,923 higher). Over the last six years the registered population has increased by 143,000 (increases of 25% in Tower Hamlets, 19% in Newham and 13% in Hackney).[22] The number of practices has remained relatively stable at approximately 145.

The majority of these 145 practices are on the same patient record system - the Egton Medical Information System (EMIS). EMIS is an electronic GP database used by 55% of English practices. Uniquely in East London the electronic records are accessible at Queen Mary University London (QMUL), pending appropriate permissions, via the Clinical Effectiveness Group (CEG), an academic unit within the medical school, co-directed by Dr. John Robson (JR) a local GP, and staffed by general practitioner academics, statisticians and facilitators. CEG's aim is to use routinely collected data from local general practices to inform needs assessment and public health planning while also contributing to the academic knowledge base in primary care and public health. It produces routine reports for local general practices and PCTs on their clinical performance. An established standing agreement enables the university-based CEG to access and audit non-identifiable service and clinical data. This central access to records made this cross-sectional analysis possible.

EMIS has three principal forms. EMIS LV is the most common and can be likened to MS DOS. It is used by 90% of practices. EMIS PCS is a newer version of EMIS used by 10% of practices. It has an interface akin to using Microsoft Windows but lacks some of the functionality of the older EMIS LV program (i.e. search and report functionality is not as

sophisticated).  EMIS web (the web based version of EMIS) is being used by a growing number of practices throughout East London (currently primarily in Newham) and is the gold standard.  It is anticipated that by the end of next year, all EMIS practices will be using EMIS web as their primary system.

Previously, before 2010, in order to run and export searches a program called MIQUEST (Morbidity Information Query and Export Syntax) was used to generate search strings and then every practice was visited in person to run that search on their individual practice. Results were then collated.  This was time intensive and costly.  With email this process was slightly speeded up.

Currently, however, for all practices using EMIS LV or EMIS PCS, all of the information is transmitted electronically via nhs.net to a server in Leeds, and reflected back to practices and authorised organisations in a read-only format in EMIS web.  The full live version of EMIS web allows practices to edit data in real time within a secure Internet cloud, as well as share with authorised organisations (including CEG).

The data are accessible to authorised organisations.  In the local area there are three main organisations with the necessary permissions to access the data (though individual practises may extend permissions to a wider group of organisations).  Firstly, data are accessible to the individual practices who can see their data only in EMIS web form, meaning they have two active forms of EMIS available to them (EMIS LV/PCS and EMIS Web).  Secondly,

the immunisation reporting data is accessed directly by the PCTs. Thirdly, the dataset is accessed by CEG who can create aggregated reports from the data, or view anonymised individual patient lines of categorical data.

CEG has an annually renewable agreement with all of the EMIS practices across East London to access patient data and use it for clinical audit and reporting purposes. Practices can freely view all of the searches and reports created by CEG and opt out of this agreement at any time. Currently, only one practice does not consent to share information.

Use of potentially identifiable clinical data requires discussion with the National Information Governance Board (NIGB) as people have not consented to their identifiable individual data being used. Typically identifiable data would include such fields as name, address, full postcode, date of birth (rather than age in years) and NHS number. In some circumstances it might also include reported cells with small numbers. CEG does not access any of this identifiable information without NIBG agreement.

CEG's anonymised reporting data takes two forms, *aggregate* reports e.g. 25% of patients have risk factors for Cardiovascular Disease (CVD), and *list* reports, with pieces of information for each individual patient e.g. age, sex, ethnicity, family history of diabetes etc.

**Selection of a diabetes risk score for East London**

Drawing on the criteria developed during the systematic review and using the shortlist of seven practical risk scores in Table 3, a risk score for use on the electronic records via CEG in East London was sought.

Using these criteria the QDScore was selected for the following reasons: [a] the intended use case was likely to be primarily in clinical consultations (although public health uses are possible through aggregation of data at CEG) and QDScore uses electronic data from GP records which are highly complete in East London, [b] the target population is ethnically and socially diverse and QDScore incorporates these variables, and [c] the expected user (the GP) in clinical practice could potentially incorporate the score automatically within the incentivised NHS Health Checks[43] (QRisk[42] is already used in this context for cardiovascular risk and therefore costs are likely to be low to add in QDScore).

The QDScore[54] gives a ten year estimate of risk of developing type 2 diabetes. The risk factors used are: age, gender, ethnicity, Townsend score of deprivation (based on Census derived unemployment, car ownership, owner occupation and overcrowding), family history of diabetes, history of cardiovascular disease, smoking status, treated hypertension, current corticosteroid usage and BMI (weight and height). It incorporates ethnicity and socio-economic deprivation together and has been created in a prospective study population

(using the QResearch database). The QResearch database is composed of a generalisable sample from England and Wales of 11 million electronic records from 551 general practices derived from EMIS.[54]

QDScore has only been derived and externally validated on populations aged 25-79 years free from diabetes at baseline.[54][72] For derivation and internal validation, two cohorts were created randomly. To derive the score 2,540,753 individuals aged 25-79 years were used. The internal validation cohort had 1,232,832 individuals. In the derivation cohort 78,081 people developed diabetes compared to 37,535 in the internal validation cohort. The study period ran for 15 years from 1993 to 2008. Ethnicity was a major risk factor as was deprivation. The AUROC was 83.4% for women and 85.3% for men. Calibration was also good as observed and predicted rates of diabetes were closely aligned. It was externally validated on the THIN database by independent researchers.[72]

The mathematical equation for QDScore is highly complex and can be found online.[107] QDScore is most accurate when electronic general practice records are complete. However, if variables are missing QDScore approximates a score using assumptions. Table 5 shows how the score handles missing values.

| Missing value | QDScore response |
| --- | --- |
| Townsend score | Assumes value is 0 (national average) |
| BMI | Substituted value used based on prediction algorithm using age, sex, ethnicity, smoking status, treated hypertension and cardiovascular disease<br><br>If BMI is out of range the processor substitutes a BMI of either 15 or 54 |
| Ethnicity | Defaults to White British |
| Smoking | Assumes not a smoker |

**Table 5: QDScore and missing values**

**Aim of the cross-sectional study**

To use a risk prediction model on the electronic records of three inner city boroughs to describe risk of diabetes to guide possible interventions for targeting groups at high risk.

The research question was:

*How feasible is it to aggregate, describe, and stratify diabetes risk in a way that meaningfully informs locality-based needs assessment and service planning?*

**Objectives**

1. To use the EMIS web database in Inner North East London to calculate 10-year diabetes risk using the QDScore on the GP-registered population aged 25-79 years.

2. To describe risk by age, gender, ethnicity, social deprivation and other co-morbidities.

3. To categorise risk and describe in detail the characteristics of the high-risk population.

The final aim and objective as presented above was initially much broader and was modified iteratively as the study progressed. An initial decision was also made to geospatially map some of the findings and this is presented separately in Section 4. However, in the course of the overall research both the cross-sectional analysis and geospatial mapping (and the latter stages of the systematic review) were conducted with

some overlap, although they have been separated in this thesis - mainly for ease of reading, and also because each has yielded separate research findings and an academic publication.

**Methods**

A cross-sectional analysis was undertaken on 519,288 electronic general practice records of all non-diabetic adults aged 25-79 years from EMIS Web across 135 out of 145 general practices in the boroughs of Tower Hamlets, Newham and City & Hackney.

Anonymised data were extracted from electronic health records in general practices using EMIS Web via N3 networks which are securely held by CEG at QMUL. The ten clinical variables needed to calculate the QDScore[54] were extracted from the records: age, gender, ethnicity, Townsend score of deprivation, family history of diabetes, personal history of cardiovascular disease, smoking status, treated hypertension, current corticosteroid usage and BMI. Additional clinical variables were also extracted for sub-group analysis including: QRisk II (referred to as QRisk throughout, and only for people over 30 years), diagnosed hypertension without treatment, estimated glomerular filtration rate (eGFR), and gestational diabetes (females only).

The QDScore was supplied as an electronic patch by the original authors (Hippisley-Cox, personal communication to RM & JR) and used to calculate risk of type 2 diabetes.[54] Basic descriptive statistics were compiled using Stata version 10.[108]

QDscore results were grouped into quintiles of type 2 diabetes risk for analysis. In some cases the bottom quintile, relating to a risk of less than 20%, was split in half for further exploration during the sub-group analysis. Certain other variables had to be categorised, and a quintiles approach was used for much of the analysis, as shown in Table 6. Risk of type 2 diabetes was further categorised as low (0-9.99% risk at ten years), medium (10-19.99%) and high (≥20%).

| Variable | Groups |
|---|---|
| Diabetes Risk | 0-19.99, 20-39.99, 40-59.99, 60-79.99, 80-100 |
| Townsend Score | -6 to 3, 4, 5, 6, 7-10 |
| BMI | 10-19.99, 20-24.99, 25-29.99, 30-34.99, 35-49.99, 50-70 |
| QRISK | 0-9, 10-19, 20-29, 30-39, 40+ |

**Table 6: Grouping techniques for variables**

Sub-group analyses using descriptive statistics, with variables used to calculate the QDScore (age, sex, ethnicity, deprivation, treated hypertension, personal history of cardiovascular disease, and body mass index), and the additional clinical variables not included in the score, were undertaken. Diagnosed hypertension with treatment was used for the QDScore patch. However, diagnosed hypertension overall with or without treatment was used for sub-group analysis. These variables were selected after discussion amongst the research team, local general practitioners and public health specialists as to which would be most useful to inform commissioning of public health interventions.

Sub-groups with higher risk were described but not unexpected as certain predictor variables within the QDScore patch lead to a higher score (e.g. certain ethnicities). Therefore statistical colinearity between the outcome (% risk) and the predictor variables rendered tests of significance misleading.

For the purposes of QDScore calculation, ethnicity codes were grouped into 17 categories based on the 2001 census. General practices in the three boroughs have access to 155 ethnic group codes. These were converted to 17 (using a standard process described with the QDScore processor manual supplied by the original authors) in order to use the QDScore electronic patch. After combining the calculated QDScore with the variables for sub-group analysis, ethnic group was reduced from the 17 categories used in the score calculation to five for ease of analysis. Ethnic categories were reduced from 17 to five as follows: White (British, Irish, other White), Black (White+Black African, White+Black Caribbean, African, Caribbean, other Black), South Asian (Bangladeshi, Pakistani, Indian, other Asian or White+Asian), Other (Chinese, other ethnic groups, other mixed groups), and Not stated or Missing (not recorded). The final category comprised: truly not stated (missing), not disclosed, or was coded at too high a level to be useful (effectively missing). Individuals who reported being of mixed Black or mixed South Asian were grouped with their parent ethnic minority group for reasons of biological plausibility.

Results were further broken down in each of the three individual boroughs so as to supply each of the relevant NHS PCTs detailed estimates of those at risk. The overall results for

all three boroughs are presented in this thesis, the detailed borough by borough results and analysis can be found in the online report.[109]

## Chapter 6: Results and Discussion

**Results**

Clinical variables for risk calculation were extracted from 135 out of 145 practices. Of the

ten practices not included in the study, one did not share data, four used non-EMIS based

systems, and five had technical problems which prevented access to data. A small number

of individual records were classified as confidential and could not be accessed. Table 7

shows the flow of data through the study.

| | Tower Hamlets | City & Hackney | Newham | Total |
|---|---|---|---|---|
| **Total number of practices** | 36 | 45 | 64 | 145 |
| **Number of practices with data available** | 35 | 40 | 60 | 135 |
| **Registered population** | 268,130 | 266,577 | 347,189 | 881,896 |
| **Aged 25-79 years** | 174,596 | 177,468 | 216,779 | 568,843 |
| **Free from Diabetes** | 163,275 | 167,685 | 199,488 | 530,448 |
| **Data available for analysis** | 163,088 | 166,762 | 189,438 | 519,288 |

**Table 7: Flow of data through cross-sectional study**

Completeness of variables that should have been routinely collected on the general practice

electronic records (n=519,288) were as follows: age (100%), gender (100%), ethnicity

(91.6%), Townsend deprivation score (99.8%), BMI (76.5%), and smoking status (96.4%).

Other variables were only recorded if positive. [Table 8] Gestational diabetes was positive

for 1.02% of females only (n=241,072).  QRisk was calculated for 53.8% of individuals

over 30 (n=410,874).

| Variable | Positive (% & number) |
|---|---|
| Family history of diabetes | 22.9   (n=119,063) |
| Personal history of cardiovascular disease | 1.9    (n=9,805) |
| Diagnosed hypertension with or without treatment | 9.3    (n=48,169) |
| Treated hypertension | 7.2    (n=37,394) |
| Current corticosteroid usage | 1.01   (n=5,240) |
| Estimated glomerular filtration rate <60 | 1.35   (n=7,026) |
| Gestational diabetes (women only) | 1.02   (n=2,466) |

**Table 8: Completeness of variable recording in cross-sectional study**

The distribution of QDScore was plotted and was determined to be heavily skewed to the

right (as shown in Figure 4).
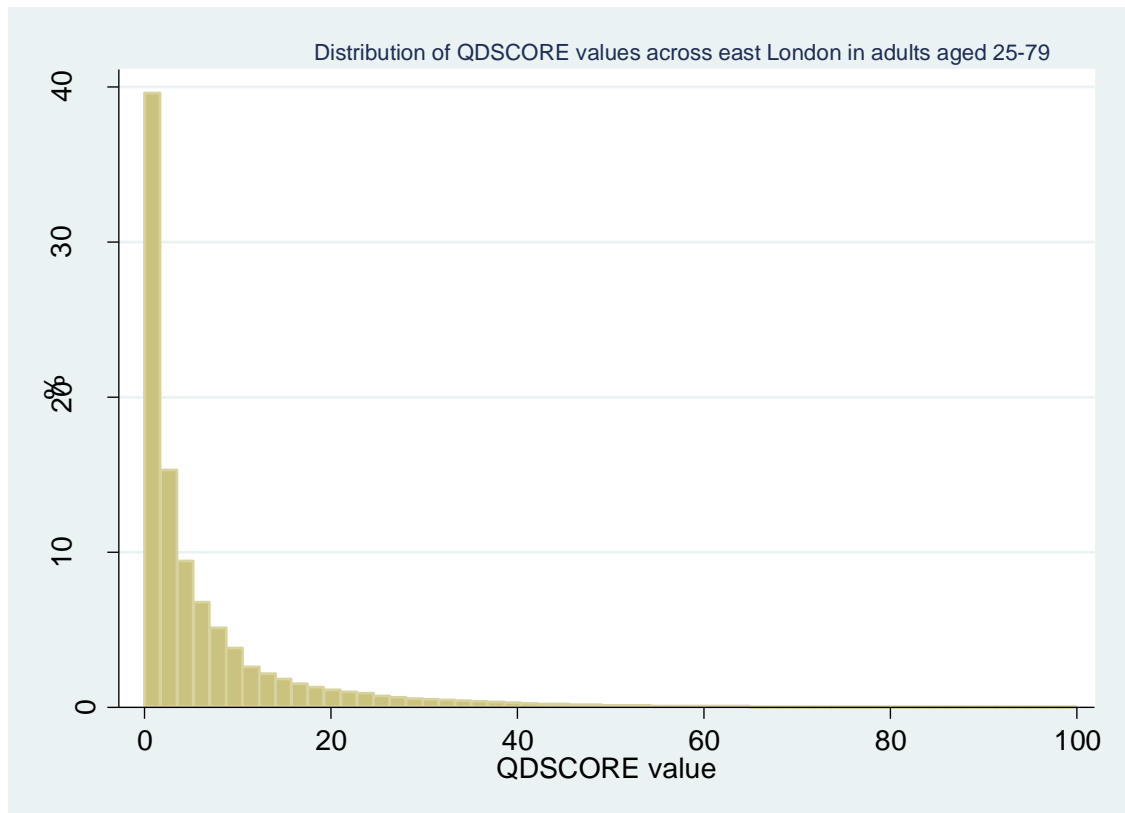
Figure 4: Distribution of QDScore


Overall, 1 in 10 people (n=51,061) in this inner-city population were at high risk (≥20%) of developing type 2 diabetes within ten years.   The risk of developing type 2 diabetes rose with age from 2.1% of 25-39 year olds (n=6,225) at high risk compared to 20.1% of 40-79 year olds (n=44,842).   Table 9 shows the proportion of individuals at low, medium and high risk of developing type 2 diabetes over the next ten years.   [Table 9, Figure 5]

| 10 year risk of developing type 2 diabetes (%) | Number of people in category | % of sample | Sex (% male) | Median score | Median age | White % (n=214,542) | South Asian % (n= 135,000) | Black % (n= 82,036) |
|---|---|---|---|---|---|---|---|---|
| 0-9.9 (low) | 410,801 | 79.1 | 53.0 | 1.8 | 34 | 83.2 | 69.4 | 72.4 |
| 10-19.9 (medium) | 57,426 | 11.1 | 55.4 | 13.8 | 49 | 9.4 | 14.2 | 15.3 |
| 20-100 (high) | 51,061 | 9.8 | 56.4 | 30.9 | 54 | 7.5 | 16.4 | 12.3 |
| Total | 519,288 | 100 | 53.6 | 2.8 | 37 | 100 | 100 | 100 |

**Table 9: Proportion of individuals at low, medium and high risk of developing type 2 diabetes over the next 10 years**
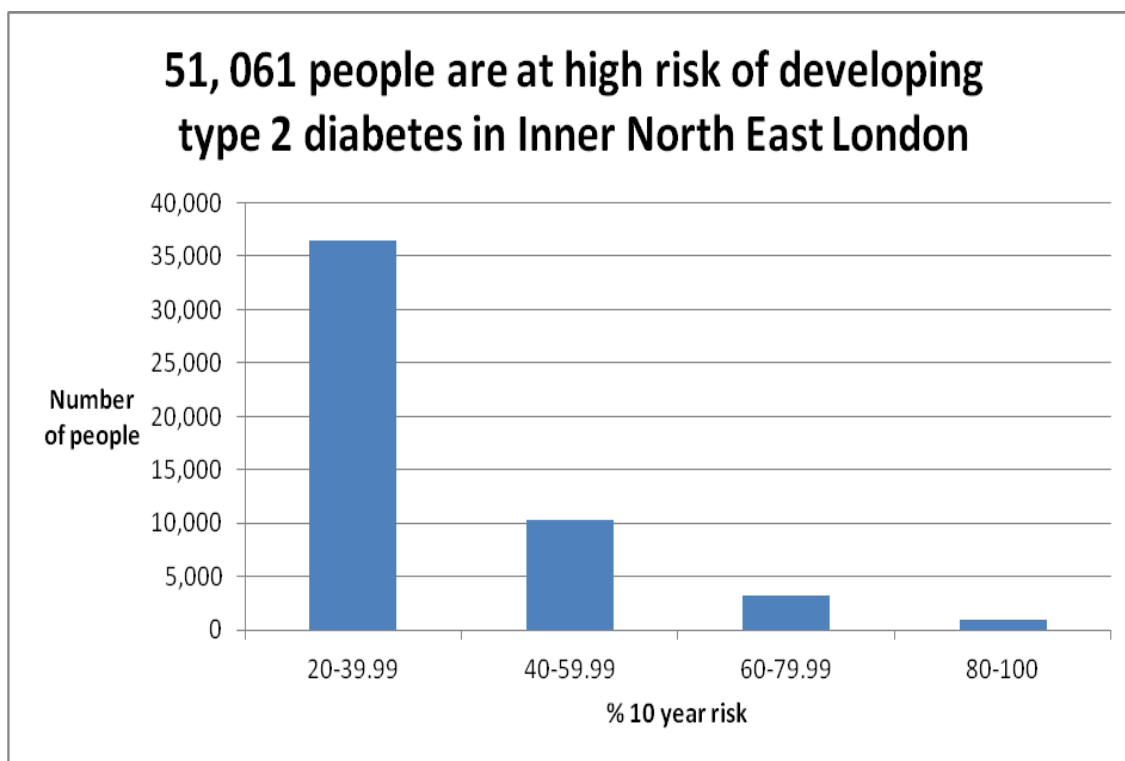


**Figure 5: Overall number of people at high risk**

The overall median QDScore varied between different ethnic groups: White 1.9%, South Asian 4.4% and Black 4.3%. More than twice as many South Asians (16.4%) were at high risk compared to the White (7.5%) population. [Figure 6]
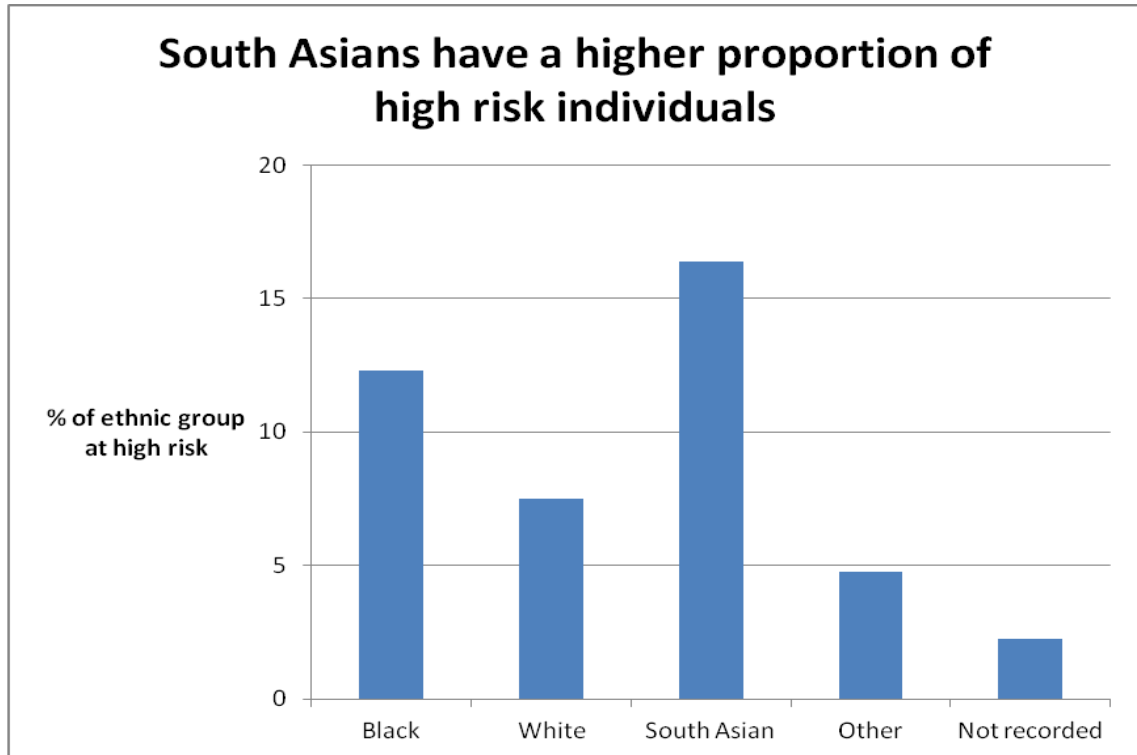


**Figure 6: Percentage of each ethnic group in the high risk category overall**

The higher the Townsend Score the more deprived the population; 15,262 people were at high risk of type 2 diabetes and in the highest band of Townsend Score. Those in the lowest band of Townsend Score (least deprived), had the lowest number of people at high risk (7.7%) compared to the highest band of Townsend Score (12.1%). South Asian ethnicity remained a strong risk factor even in non-deprived sub-populations. The most affluent South Asians (Townsend score -6 to 3) had a higher proportion at high risk than

the most deprived at high risk from all ethnic groups, showing the impact of ethnicity on

risk.  [Figure 7]

<figure>
**As Townsend Score increases so does QDscore**

$R^2 = 0.9846$

$R^2 = 0.9398$

Legend:
- % of group at high risk
- % of South Asians at high risk

X-axis (Townsend Score): -6 to 3, 4, 5, 6, 7 to 10
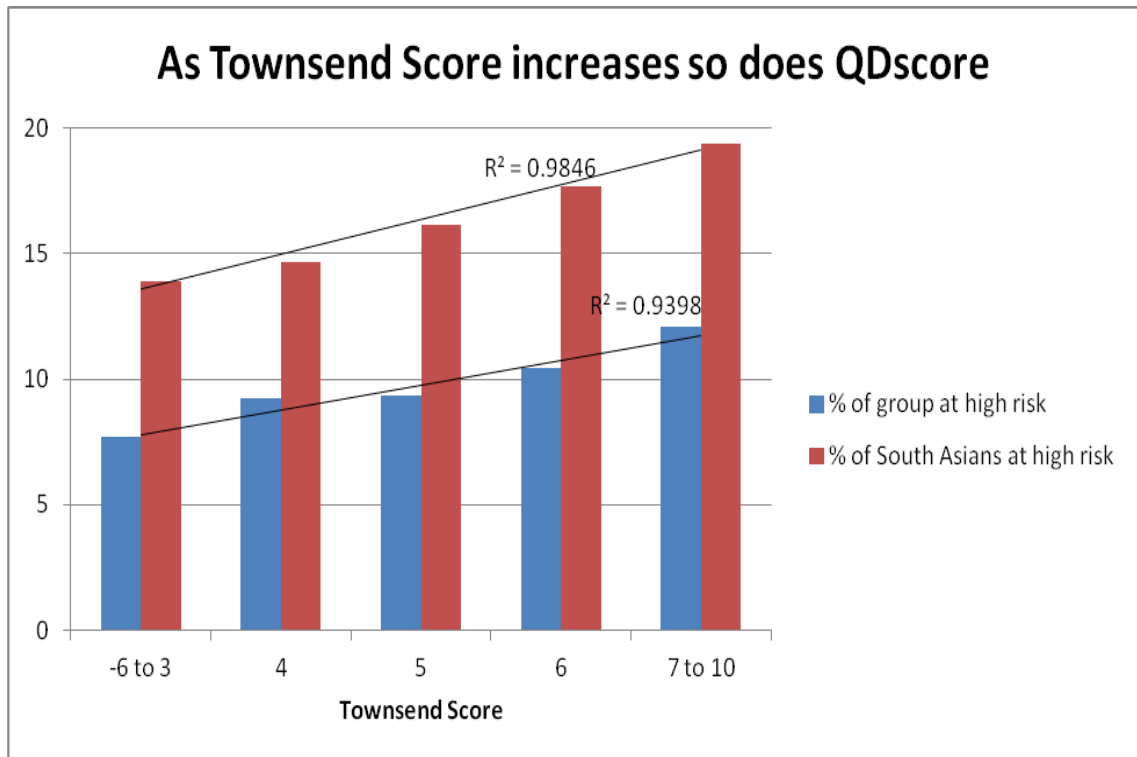</figure>

**Figure 7: Percentage of adults at high risk of developing type 2 diabetes in each Townsend Score band**

Cardiovascular risk as estimated by QRisk was closely associated with high risk of type 2

diabetes.  For QRisk 0-9, 9.7% (15,516) were at high risk for type 2 diabetes, compared to

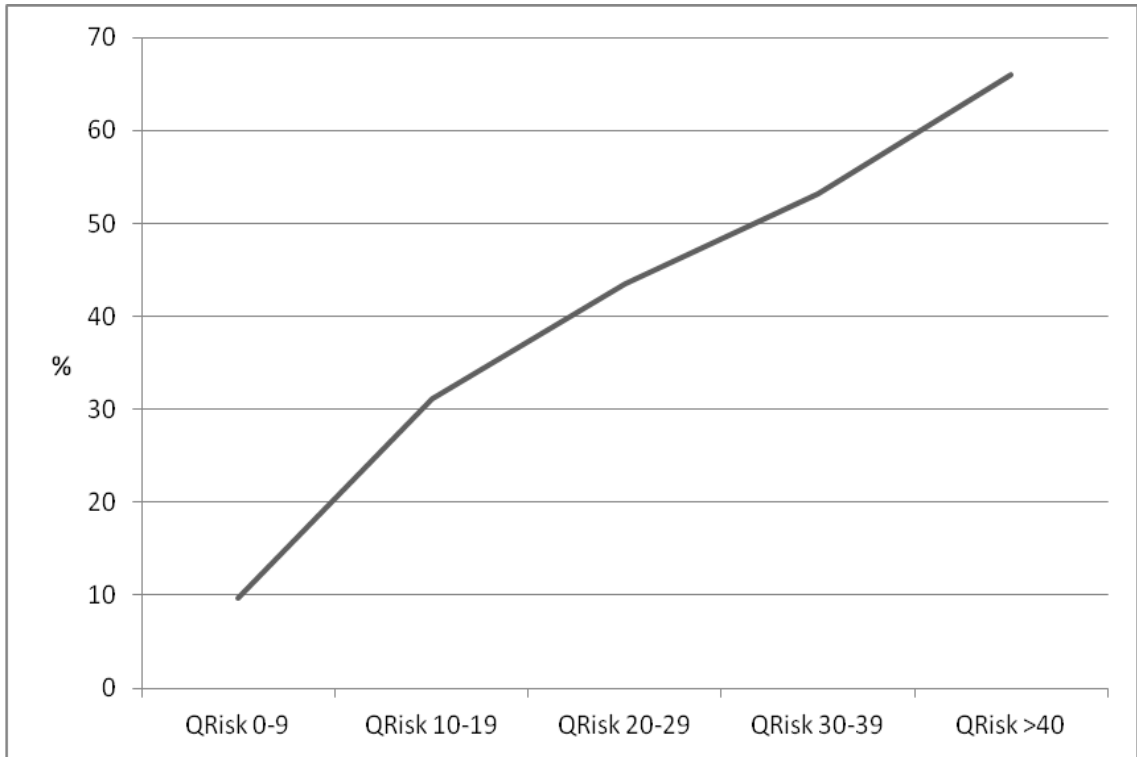31.1% (12,487) for QRisk 10-19, and 47.7% (9,839) for QRisk ≥20.  [Figure 8]

**Figure 8: Percentage of adults over 30 at high risk of developing type 2 diabetes in each QRisk band**


Similarly as shown in Table 10 vascular co-morbidity, eGFR (<60mmls/min/1.73$^2$, chronic kidney disease stage 3 or greater), gestational diabetes and increasing BMI all increased the chance of being at high risk of developing type 2 diabetes.

| | Number of people at high risk | Number of people in category | Percentage at high risk |
|---|---|---|---|
| **CVD (IHD/stroke/TIA)** | 5,637 | 9,864* | 57.1 |
| **Hypertension (diagnosed+/-treatment)** | 23,102 | 48,169** | 48.0 |
| **eGFR<60** | 2,905 | 7,026 | 41.3 |
| **Gestational diabetes** | 446 | 2,466 | 18.1 |
| **BMI>30** | 32,564 | 76,162 | 42.8 |
| **QRisk≥20** | 9,839 | 20,629 | 47.7 |

**Table 10: Percentage of people at high risk in various co-morbidity groups**

*\*9805 of 9864 had the diagnosis included in the QDScore calculator as nine codes were added later for sub-group analysis which should have been included in the original score calculation.*
*\*\*37, 394 had diagnosed and treated hypertension and this diagnosis was included in the QDscore calculation.*

In Newham 1 in 8 people (22,513) were at high risk of type 2 diabetes overall (≥20%), compared to 1 in 11 people (15,304) in Tower Hamlets and 1 in 13 people (13,244) in City and Hackney. Other small differences were present between each of the boroughs, but these are not described in detail in this thesis.

**Discussion**

The QDScore was successfully used on half a million electronic primary care records to describe the socially patterned risk of developing type 2 diabetes for an entire inner city population. Risk of adults developing type 2 diabetes was universally high: 1 in 10 people (51, 061) were at high risk ($\geq$20%) of developing type 2 diabetes within ten years. In Newham 1 in 8 people (22, 513) were at high risk, compared to 1 in 11 people (15, 304) in Tower Hamlets and 1 in 13 people (13, 244) in City and Hackney.

Increasing age and male sex conferred additional risk. The median age of people at low risk (<10%) was 34 years compared to 49 years for those at medium risk (10-20%) and 54 years for those at high risk ($\geq$20%). Of 25-39 year olds 2.1% were at high risk compared to 20.1% of 40-79 year olds.

It is well established that social and ethnic diversity of populations heavily influence chronic disease risk. In the cross-sectional study ethnicity and risk of diabetes were closely associated. More than twice as many South Asians (16.4%) were at high risk compared to the White (7.5%) population. However, high risk was not confined to South Asians. In Newham 10.5% of the White population was at high risk ($\geq$20%) compared to 6.5% in Tower Hamlets and 6.5% in City & Hackney.

Socio-economic deprivation was associated with increased risk and ethnicity increased this association. Those in the lowest band of Townsend Score have the lowest proportion of people (7.7%) who are at high risk (≥20%) compared to the highest band of Townsend Score (12.1%). Within the South Asian population at high risk (22,126) the proportion of people at high risk increased and was higher than the overall population in every Townsend Score band.

Obesity and cardiovascular co-morbidity substantially increased risk of developing type 2 diabetes. For example, 76,162 people in the cohort had a BMI greater than 30, and 42.8% of these were high risk for developing type 2 diabetes. There were 57.2% of people with cardiovascular disease, 48.0% with hypertension, 41.3% with chronic kidney disease, and 18.1% with gestational diabetes who were at high risk for developing type 2 diabetes, compared to 9.8% of the cohort overall.

The overlap with QRisk was extensive. Of those with a QRisk score ≥20 (meaning, ≥20% risk of developing cardiovascular disease in 10 years), 48% were also at high risk for developing type 2 diabetes. This underlines the need to combine preventive interventions for these common conditions with overlapping risk factors.

The extent of these findings locally was quantified and described. Although results are not surprising, especially as QDScore assigns higher values to known risk factors, detailed population sub-group analysis have high potential to inform targeted interventions.

The dataset had high completeness due to previous investment and long standing relationships between the university, general practices and the primary care trusts. This demonstrates the type of population statistics that can be generated using a risk prediction model on electronic records, and the rich level of detail which sub-group analysis can generate. For example, knowing the percentage of an ethnic group at high risk has the potential to inform targeted preventive measures through social marketing.

The QDScore has only been validated to estimate risk of diabetes for individuals aged 25-79 years. A large proportion of those registered with a GP (n=313,053; 35.5%) was outside this age range, reflecting the young population in East London. Type 2 diabetes is increasingly common in younger age groups making this a weakness of the cross-sectional study.

Colinearity between outcome and predictor variables prevented tests of statistical significance, but as the principle purpose of the descriptive statistics is for service planning, and trends were very clear (e.g. association with ethnicity), this is unlikely to impact on the routine use of such data.

Whilst many studies exist describing diabetes risk models and scores[94 96 110], relatively few validated scores have described 'real world' applications. The Finnish Type 2 Diabetes Risk Score (FINDRISC) was used in a survey of 400 adults aged 20 to 73 years in Libya; approximately 12% were at high or very high risk of developing diabetes over ten years.[111]

Both FINDRISC and the Indian Diabetes Risk Score were used on 198 migrants in Norway identified through mosques, Norwegian classes, and directly in shops and on the street; with FINDRISC 29% were at great/extreme risk (the two highest categories) over ten years.[112]   More recently QRISK 2, Framingham and Joint British Societies' 2 (cardiovascular risk scoring systems) were offered to 434 South Asians attending Hindu temples; using QRISK 2, 15% had a risk greater than 20% of developing cardiovascular disease over ten years.[113]  However, all these studies were small and exploratory, and were focused mainly on identifying individuals rather than managing risk in entire populations – like that presented from this cross-sectional study.

Yet, despite it being feasible to extract a vast quantity of anonymised data to determine population estimates of risk, and perform detailed subgroup analysis, making the targeting of high risk groups possible, I was concerned about the accessibility of this health information to local commissioners.  There were several reasons for these concerns.  First, although a summary of the main findings are presented in this thesis, in the report to the local NHS partners[109] pages of detailed figures, tables, and bullet pointed health information findings were presented borough by borough.  Having worked in both the Department of Health and in a PCT, I was used to observing officials, managers and doctors spending only moments glancing over such data, usually in a time-limited allocated slot in the midst of vast swathes of other business e.g. as a 15 minute agenda item during a two hour monthly vascular care quality group meeting.  Despite their best intentions, it appears that many professionals in these situations only engage with the material when they

are actually at the meeting, as oppose to studying papers in advance and perhaps making better use of the meeting time to critically appraise findings and decide on change in policy. For this reason, presenting data in as compelling a way as possible is essential.

Second, the NHS is undergoing radical restructuring, including the transfer of commissioning of clinical and public health services to clinical commissioning groups (comprising general practitioners, secondary care clinicians, nurses and lay members).[114] Part of the public health function will shortly be transferred to local authorities, who are traditionally responsible for urban planning and environmental health. Commissioning and local authority bodies will therefore need health information in an easily accessible format in order to plan, procure, monitor, evaluate and coordinate clinical and public health interventions and neighbourhood initiatives. My opinion is that tables of data and figures alone will not enable these new decision makers to commission effectively.

Third, in the UK, information for health planning and management is ubiquitous.[115] This has been handled traditionally by public health specialists and ancillary staff. Members of the new clinical commissioning groups will need to possess skills in handling health information in order to commission services.[116] Yet, some evidence suggests, at least for general practitioners, that both skills in handling and using health information for commissioning may be limited.[117] This could be linked to: [a] lack of training in handling and processing population level data, [b] lack of skills in prioritising health information based on health needs as oppose to exclusively service demands or cost savings, and [c]

lack of experience in using health information selectively to plan and manage services and public health interventions. Key to addressing this, in addition to training, is the presentation of health information in easily accessible formats, which facilitates clinical commissioning groups to develop expertise in using health information for health planning and management.

Whilst these concerns may have some merit, there is no 'one option fits all' approach to health information presentation. And, although not explored in detail in this thesis, the style of the presenter is likely as important as the materials to hand. Nevertheless, I was interested in further exploring geospatial mapping of selected results from the cross-sectional study, first to explore whether it was possible (as it would potentially entail data extraction of half a million postcodes), and second to begin to assess whether this could aid operationalising the recently-published NICE guidance on diabetes prevention, which recommends the use of: 'local and national tools … to identify local communities at high risk of developing diabetes to assess their specific needs.'[13]

## SECTION 4: Geospatial Mapping

## Chapter 7: Background, Aims, Methods

In this section I describe a geospatial study using some of the data from the cross-sectional review described in Section 3, and new data containing a geospatial locator for each electronic patient record. This was published in BMJ Open in February 2012.[118] [Appendix 6] The extraction of postcodes for geospatial mapping raised a number of information governance issues and these were separately explored in a specific information governance paper currently submitted to BMC Public Health. [Appendix 7] Overall the study attracted significant media interest including from the British Broadcasting Corporation (BBC) and the Daily Mail. In Chapter 7 I outline background, aims and methods, and in Chapter 8 I present the geospatial maps with accompanying discussion. The discussion includes two sub-sections considering information governance issues, and the media reporting of the findings.

**Background**

Historically, mapping has often been used in a public health context. Early pioneers of geospatial mapping of health information included Dr. Alfred Haviland who published in 1892 his: 'Geographical distribution of diseases in Great Britain,' and most famously the mapping by Dr. John Snow of cholera cases surrounding a water pump in Broad Street, London, in the mid 1800s.[119] [120]

Using health information in the format of geospatial maps enables exploration of spatial patterns and geographical associations with wider social determinants of health.[121][122]  For example, the International Geographical Union presented a report in Washington, USA in 1952 from its Commission on: 'The study of geographical factors concerned with cause and effect in health and disease.'[123]

Presenting health information as maps has various uses, including: [a] idea generation and theory formation at project initiation, [b] during scrutiny of research results, and [c] to assist with visual presentation of findings to relevant stakeholders.[122]

Geovisualisation – the use of computer-aided graphical methods (Geographic Information Systems - GIS) to visualise geospatial information[124] – is a technique which has begun to be used to help guide health service planning, public health interventions and inform the public about disease 'hot spots'.[125]  A well-known use of this technique are the maps of obesity produced by the Center for Disease Control in the USA, which have shown higher prevalence in the southern states and a shift in prevalence from low (shown in blue) through high (shown in red) over the past 40 years.[126]

Geospatial mapping of self-reported questionnaire data has shown the USA to have a 'diabetes belt' (i.e. a band of states with high prevalence of this condition) in the south-east of the country linked to distribution of the known risk factors of obesity, inactivity, and African-American ethnicity.[127]  Small-area geographical variation in diabetes prevalence

110

has also been mapped in a single city in Canada using research survey data, and links demonstrated with the geographical distribution of social and environmental determinants including family income, education, aboriginal status and neighbourhood crime.[128]  In the UK, small-area mapping of coronary heart disease morbidity and mortality using multiple data sources (e.g. hospital admission statistics and mortality statistics) has been linked to social and environmental risk factors (e.g. income and ethnicity) and geographical 'hot spots' of coronary heart disease demonstrated in localities where these risk factors are clustered.[129]  Data from a UK population-based register of arthritis has been used to identify geographical clusters of polyarthritis.[130]

A key aspect of rigour in geovisualisation of disease or risk of disease is the completeness, accuracy, timeliness, accessibility and granularity of the primary data from which the maps are constructed, and in particular the extent to which the data are capable of illuminating the fine-grained geographical variability needed to inform locality-based health or environmental interventions.

Unlike USA and Canada, the UK has the advantage of near-universal registration with general practitioners, whose records are at an advanced state of computerisation.[131] Quality of electronically held data is high in most practices, partly due to the national financial incentive scheme for general practice, the Quality and Outcomes Framework, a component of which is chronic disease management.[132]  Aggregated data from Quality and Outcomes Framework returns has been used to model estimates of disease prevalence by locality.[133]

111

Overall using geospatial maps for chronic disease risk (such as type 2 diabetes) at small area level in local districts is at an embryological stage. Demand is likely to increase as more lay practitioners without epidemiological training adopt health planning roles. This is even more likely given the recent advances in general practice computer systems including the remote server 'cloud' storage of records, with staff gaining access via the World Wide Web rather than records held on practice based servers. This allows authorised staff to undertake complex data searches across large numbers of practices, allowing the possible use of local general practice records to be used as the data source for sophisticated mapping of disease or risk factors by small geographical area. However, accessing and using personal medical data for this purpose raises significant practical, technical, ethical and information governance challenges.

Small area geospatial mapping of disease risk factors using electronic primary care records as the data source and oriented primarily to an audience of local health planners is important when considering dense urban areas where a street may separate relatively poor and affluent neighbourhoods. Models estimating disease prevalence often show greatest discrepancy between observed and expected prevalence in areas of social complexity, suggesting that small-area mapping may be particularly useful in such areas.[134] As well as these potential uses it also has the possibility of improving translatability of health information for new commissioners.

The geospatial study was based in Tower Hamlets only.  Geographically Tower Hamlets is a well known inner-city district in the East End of London UK, known internationally for its vibrant street life, restaurants and culture, and also for its socio-economic deprivation and poor health outcomes.  Tower Hamlets is home to a large British-Bangladeshi population and to more recent migrants from Africa and to a white British working class population.  The borough includes significant pockets of deprivation, mainly in high-rise estates, alongside pockets of affluence such as riverside suburbs in the South and parkside ones in the North.  Tower Hamlets thus exemplifies the challenges facing providers and commissioners planning for culturally diverse and disadvantaged populations in inner city urban areas.

**Aim of the geospatial study**

To explore the feasibility of producing small area geospatial maps of chronic disease risk for use by clinical commissioning groups and public health teams.

The research question was:

*Is it feasible to map geospatially an entire population's risk of type 2 diabetes in a way that could lead to engagement of commissioners on the usefulness and applicability of the findings?*

**Objectives**

1. To map the percentage of people at high risk of type 2 diabetes for each Lower Super Output Area in Tower Hamlets.

2. To compare geospatially the percentage of people at high risk with deprivation, ethnicity and selected social and environmental determinants of health.

3. To trial several different mapping methods.

4. To compare modern maps of disease risk with historical maps of poverty.

5. To assess the feasibility of producing maps.

6. To consider the extent to which such information would be useful to clinical commissioners and local authorities engaged in neighbourhood regeneration.

Initially I had hoped that geospatial mapping would be possible for all 519,288 electronic records in the cross-sectional study. However, it transpired that due to complex data sharing agreements between general practices, PCTs and CEG, individual patient postcodes were only available to be downloaded from general practices in Tower Hamlets. This was a disappointment as having worked in the region I was aware that Tower Hamlets was often perceived as the area with the most investment. To help rectify this imbalance towards Tower Hamlets I was able to produce one map at general practice (rather than patient) level which covered the whole region and this was used as the *flagship* map on the front page of the CEG special report.[109] The focus on Tower Hamlets did, however, allow separation of the cross-sectional study from the geospatial study and has resulted in a more in-depth analysis of small area mapping. The aims and objectives were refined as this part of the research progressed. For publication in BMJ Open, the findings were framed within the paradigm of chronic disease risk to make the results more widely applicable, and to build on the United Nations summit on chronic diseases which took place during the course of the research. A particular focus of the study was to identify the practicalities and information governance hurdles around the secondary uses of general practice data at a time when local general practice led commissioning groups were being established. It transpired that there were a number of information governance hurdles, which are discussed in Chapter 8.

**Methods**

Two complementary data sources were used: postcode with clinical risk factors for individual residents of Tower Hamlets, drawn from primary care electronic records; and social and environmental determinants of health, drawn from local authority registers and nationally available data at lower super output area level (relating to around 400 households/1,000-1,500 people) or middle super output area (around 2,000 households/5,000-7,200 people).

Using the electronic general practice record system, a cohort was identified comprising all non-diabetic individuals aged 25-79 years in Tower Hamlets from 35 out of 36 general practices that used the same computer system. Data download was carried out on the CEG secure N3 networks (which only authorised third parties and NHS organisations can use). In order to overcome the information governance hurdle of preventing postcode linking to clinical variables it was necessary to first download clinical variables attached to a pseudonymised identifier (n=163,275 – 'dataset 1'). And then, postcode was downloaded separately attached to the same pseudonymised identifier (n=159,353 – 'dataset 2'). The reduction in numbers was due to two practices that could not share postcode for technical reasons. We converted Tower Hamlets postcode districts (n = 8,911) to lower super output area (n=130) using an electronic lookup table.[135] Dataset 2 (with lower superoutput area, but without postcode) was linked using the pseudonymised identifier to dataset 1. Thus, each individual record in the final dataset comprised a set of individual-level clinical risk

factors plus a lower super output area level indicator of geographical locality which could be related to local and nationally available statistics.

The local authority dataset, extracted at middle super output area, comprised: [a] fast food outlets per capita (n=371), [b] green spaces per square kilometre, and [c] population density per square kilometre. Fast food outlets were identified using local authority registry data for codes X15 'takeaway' and X17 'restaurants'. All X17 codes were manually reviewed by Dr Dianna Smith (DS) and I, and premises unlikely to serve fast food as a major part of their business based on their registration details were removed. This step was necessary because large corporate fast food chains such as McDonalds were registered as 'restaurants' rather than 'takeaways'. Green spaces were quantified at the lower super output area level using the Generalised Land Use Database from 2005, which provides data on the area (in square kilometres) in each lower super output area dedicated to public green space. This did not include private gardens.[136] Population density was defined as the total population size of the middle super output area divided by the area in square kilometres. This was calculated from the Office for National Statistics (ONS) mid-year population estimates for 2010, the most recent available.

For each individual in the final dataset, 10-year risk of diabetes was estimated using the QDScore.[54]

There was no previous methodology that could be found for describing how chronic disease risk from an entire borough's set of primary care electronic records should be displayed by lower super output area. Methodological principles were therefore applied from other relevant research.[122] Determining how to display and group data, such as using deciles versus quintiles or percentage at risk versus median risk score (as QDScore was not normally distributed as described in Chapter 6) required consultation and consensus-building with relevant local partners including: academics, general practitioners and the director of public health. The final selection of display formats reflected what these consultees considered the most meaningful framings of the data.

Three different geospatial mapping techniques were employed using ArcGIS version 9.2[137] and Adobe Illustrator version 10. In the 'basic' (choropleth style) maps the high-risk (10 year risk of ≥20%) population was displayed by lower super output area as a proportion of the denominator (non-diabetic adults aged 25-79 years). A basic map was also created of the Index of Multiple Deprivation score 2010[106] to allow a visual comparison between high risk of type 2 diabetes and a different indicator of deprivation than that used within the QDScore. Statistical analysis of correlation was not performed due to an unquantified degree of colinearity between Townsend score which is used in the QDScore and Index of Multiple Deprivation. Basic maps thus presented the data as geographically defined lower super output areas (typically defined by street blocks) in different shades of colour. A list of GP practices and hospitals were located using their postcode. They were located in GIS

using the centre of each postcode. This analysis was performed to demonstrate the potential usefulness of informing local practice geographical needs assessment.

The 'heat maps' assigned the proportion at high risk to the population-weighted centroid for each lower super output area. A Kriging procedure was used to create an interpolated surface of risk.[138] Kriging estimates the value of risk between data points where the value of risk is known. In lay terms, it uses all the values for each small area on a map and estimates values between these points. In effect this creates more data points on a map and allows finer detail to be plotted. One use of the Kriging procedure is to create a heat map which shows a gradation of risk from low to high along a spectrum of colours. Heat maps offer a statistically 'smoothed' presentation of data in which the lower super output area blocks are no longer visible. One heat map for all three PCT areas was produced using data at the level of general practices. For this, the EMIS code of each general practice was used to identify all registered people aged 25-79 years at high risk of diabetes as the numerator, and all people aged 25-79 years without diabetes as the denominator, therefore calculating a proportion at high risk for each general practice. This enabled geospatially mapping high risk of diabetes across a larger area, including Tower Hamlets, Newham and City & Hackney. In total 519,288 records were used for this one map across 135/145 practices. The postcode of each general practice was used with a recent ONS postcode look-up table (August 2010) which identified an exact location in space for the general practice with a grid reference. One general practice's postcode did not have a grid reference as it was a new build. For this practice we used an adjacent postcode to locate a grid reference.

119

Several practices (n=20) had the same postcode. For these the final digit of the x + y co-ordinate was changed by 1 so that they could be separated in space by approximately 3 metres. Proportions of high risk individuals per practice (n=135) was mapped using a Kriging procedure.

The 'ring maps' are a relatively new technique which allows factors of interest (such as putative environmental determinants) to be displayed circumferentially around a map.[121] To produce these, data was aggregated to the level of middle super output area (n=31) and presented as quintiles of risk. The following data were assembled for each middle super output area: [a] fast food outlets per capita, [b] percentage of non-green space, and [c] population density per square kilometre. Using a validated adjustment procedure,[121] each was divided into highest quartile, middle 50% (2nd and 3rd quartiles), and lowest quartile. The ring map thus gives a less granular picture of the geographical distribution of a variable but allows additional mapping of factors that might influence this variable in each locality. A second ring map displaying South Asian ethnicity and unemployment score from the IMD was also created using a similar technique. This was a less valid method because of overlap and colinearity with the variables of the QDScore, but was created during exploration of this method of mapping.

A final analysis was undertaken using the maps of high risk of diabetes and deprivation, and compared to Charles Booth's historical maps of poverty in the East End of London.[139] Given that significant urban change had occurred since the 114 years between the historical

and modern maps, and differences in scales and methods, this simple analysis was limited to visual inspection of the maps and circling of areas with persistent health risks.

The whole exploratory geospatial study was made possible by a number of key partnerships between QMUL, general practices and the PCT. These were similar to those for the cross-sectional study described in Section 3. For the geospatial study there was a new link to the Department of Geography at QMUL.

Because of the extraction of postcodes specific permissions were requested for this part of the research. The study was classed as service 'audit' and deemed outwith its remit by the local NHS Research Ethics Committee. The local information governance group representing the general practices at the PCT agreed to the study, and advice on data handling and mapping was also sought from the NIGB.

The tasks of identifying, extracting, manipulating, sharing, summarising and presenting the data, presented complex practical, technical and information governance challenges. To capture these, a dataset was collected comprising documents (protocols, service level agreements, agendas and minutes of meetings), and correspondence (letters, emails, notes of telephone calls). Those represented in this dataset included the NHS Research and Ethics Board, University Departments, Tower Hamlets PCT, local general practitioners and public health specialists, and the NIGB.

This dataset was analyzed by applying a theoretical framework developed previously to study the complex organisational, social and political issues involved in introducing a nationally shared electronic medical record.[140] Specifically considered was: [a] information governance challenges, [b] practical challenges, such as the ease with which procedures could actually be carried out, and [c] technical challenges including issues of data security, downloading and interoperability.

# Chapter 8: Results and Discussion

## Results

Completeness of general practice records in the selected cohort aged 25-79 years without diabetes that should have been routinely collected (dataset 1, n=163,275 − excluding 187 where patient permission was withheld) was as follows: age (100%), gender (100%), ethnicity (92.1%), Townsend deprivation score (99.7%), BMI (76.4%), and smoking status (96.3%). Other variables were only recorded if positive. [Table 11]

| Variable | Positive (% & number) |
|---|---|
| Family history of diabetes | 21.5   (n=127,995) |
| Personal history of cardiovascular disease | 1.8   (n=2,972) |
| Treated hypertension | 5.1   (n=8,244) |
| Current corticosteroid usage | 0.5   (n=833) |

**Table 11: Completeness of variable recording in geospatial study**

Records could not be generated or were removed if: [a] The general practice was not able to share the data for technical reasons (n=3,922) or patient permission was withheld (n=187), [b] the individual record contained no postcode (n=29) or lower super output area was not calculable from the available postcode (n=275), [c] the geographic location was outside Tower Hamlets (n=1,813), or [d] there was a mismatch between records in set 1 and set 2 (n=4). This left 157,045 records for analysis (96.2%) representing 33 out of 36 general practices.
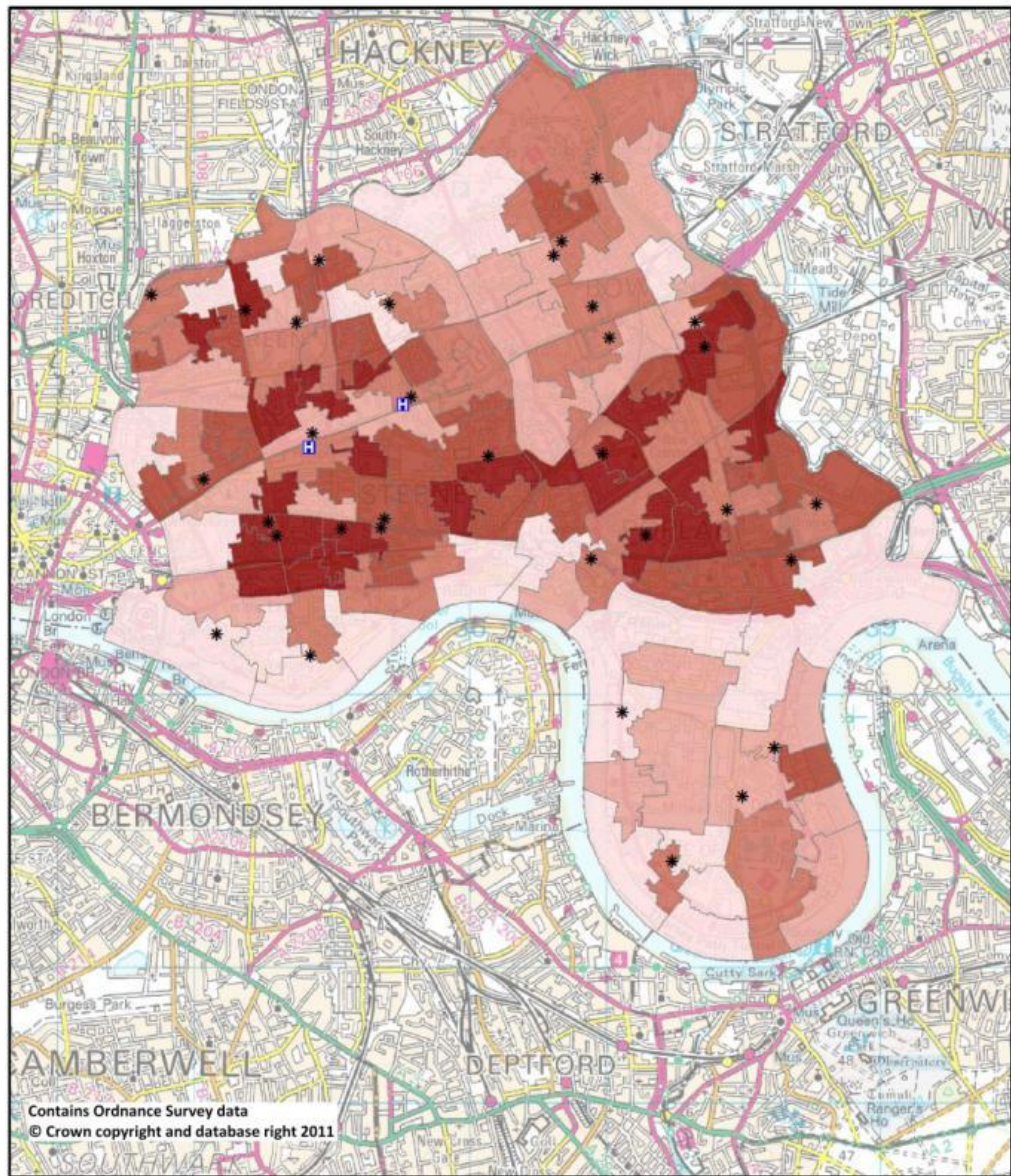
Reducing the list of restaurants to those with a major business purpose of takeaway food resulted in a total sample of 371 outlets, shown in Table 12 below.

| Reason removed | 895 (Codes X15 + X17) |
|---|---|
| Removed no postcode | 62 |
| Removed as staff restaurant, kitchen or canteen | 142 |
| Removed as usage unclear | 74 |
| Removed as Cafe | 149 |
| Removed as Bar | 8 |
| Removed as Restaurant | 88 |
| Removed as closed | 1 |
| Final included | **371** |

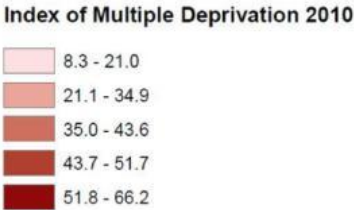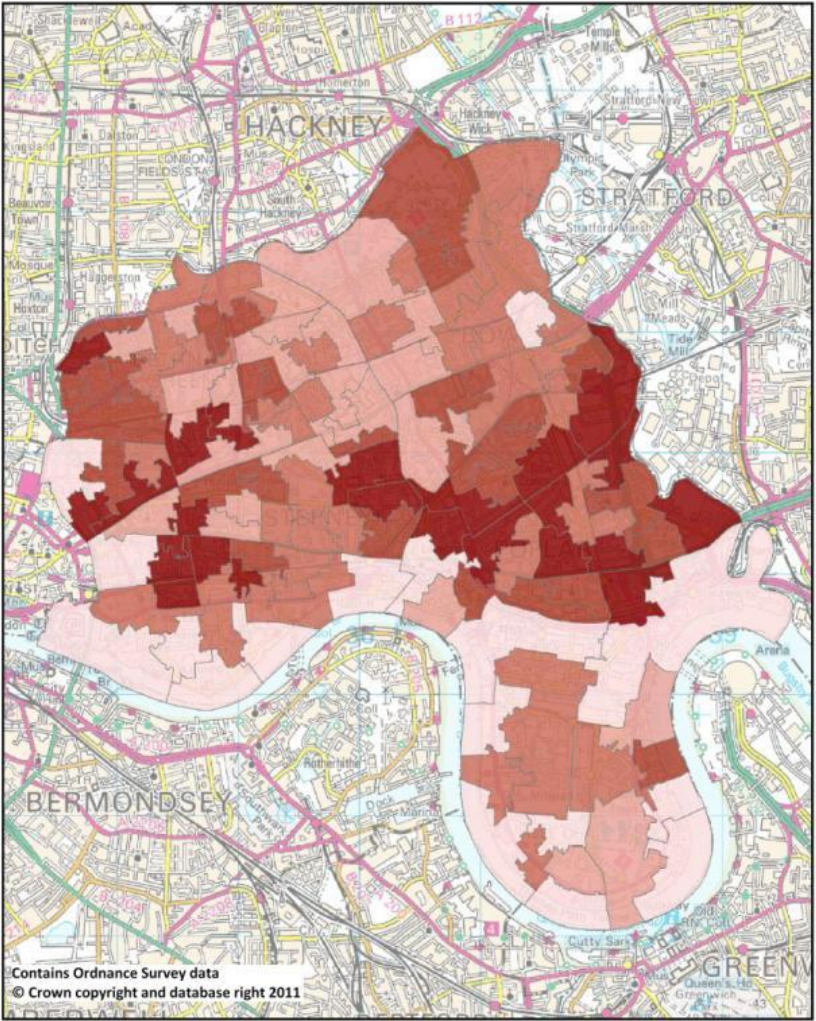**Table 12: Fast food restaurants flow of data**

Of the data which was used in the mapping (n=157,045) 9.48% of people (n=14,885) were at high risk of developing type 2 diabetes within ten years.

The *basic* Map 1 illustrates the variation in prevalence of high diabetes risk across lower super output areas in Tower Hamlets, with a maximum of 17.3% of the non-diabetic population being at high risk. General practices and hospitals are also shown. The areas of highest prevalence for diabetes risk were distributed on either side of the main east-west road (the A11) which transects the borough and corresponds with well-known deprived housing estates and high-rise blocks of flats on either side of this road.

**Percentage at high risk**

| | |
|---|---|
| 0.0 - 5.4 | **H** Hospitals |
| 5.5 - 8.3 | * GP practices |
| 8.4 - 11.3 | |
| 11.4 - 13.7 | |
| 13.8 - 17.3 | |

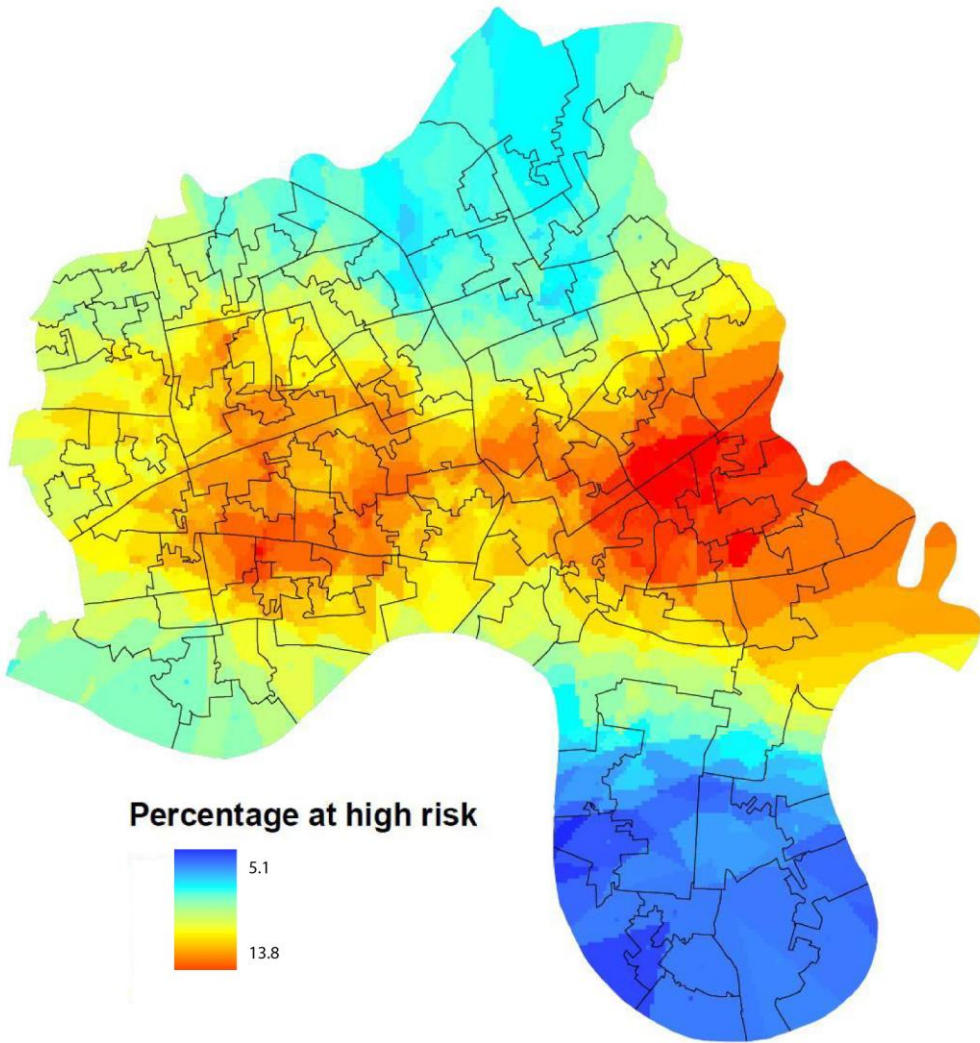**Map 1: Percentage at high risk of type 2 diabetes by lower super output area**

The *basic* Map 2 of IMD scores by lower super output area showed a near-identical geographical distribution with high diabetes risk.



**Index of Multiple Deprivation 2010**

| | |
|---|---|
| | 8.3 - 21.0 |
| | 21.1 - 34.9 |
| | 35.0 - 43.6 |
| | 43.7 - 51.7 |
| | 51.8 - 66.2 |

**Map 2: Index of Multiple Deprivation 2010 by lower super output area**

The *heat* Map 3 shows the same information as in Map 1 but displayed as a globally smoothed surface over the entire geographic area. The prevalence of high diabetes risk in this smoothed version of the data varied from 5.1 to 13.8%.
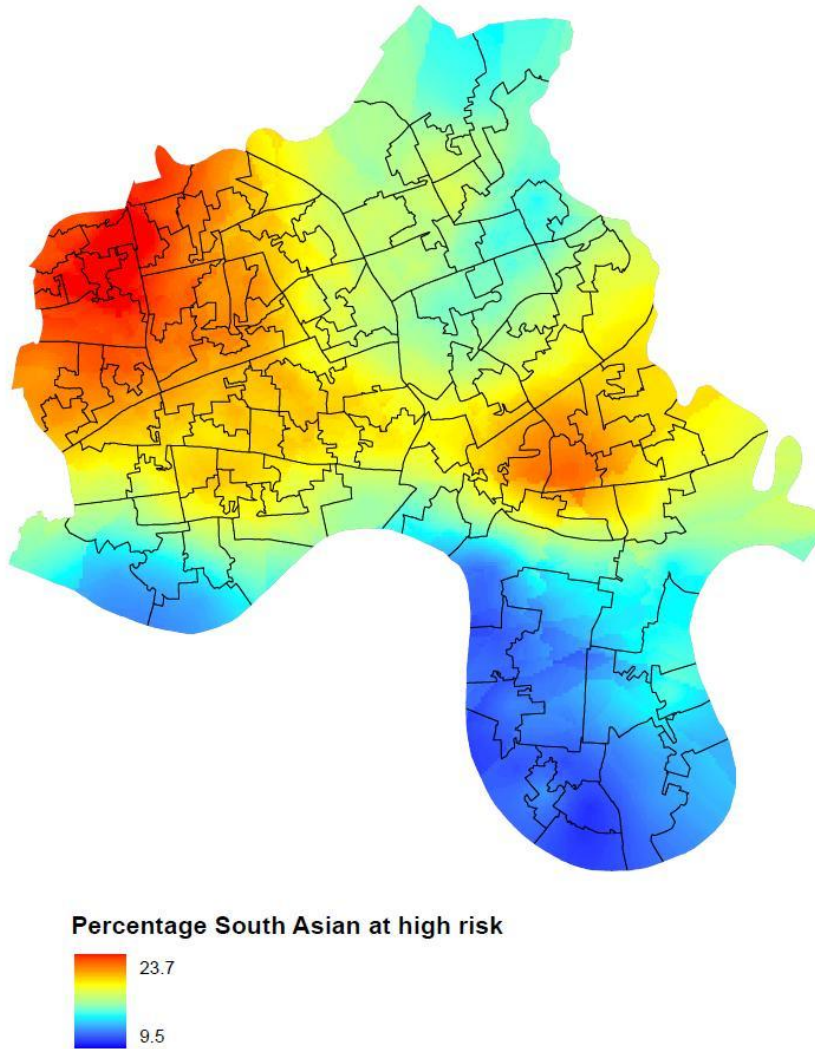


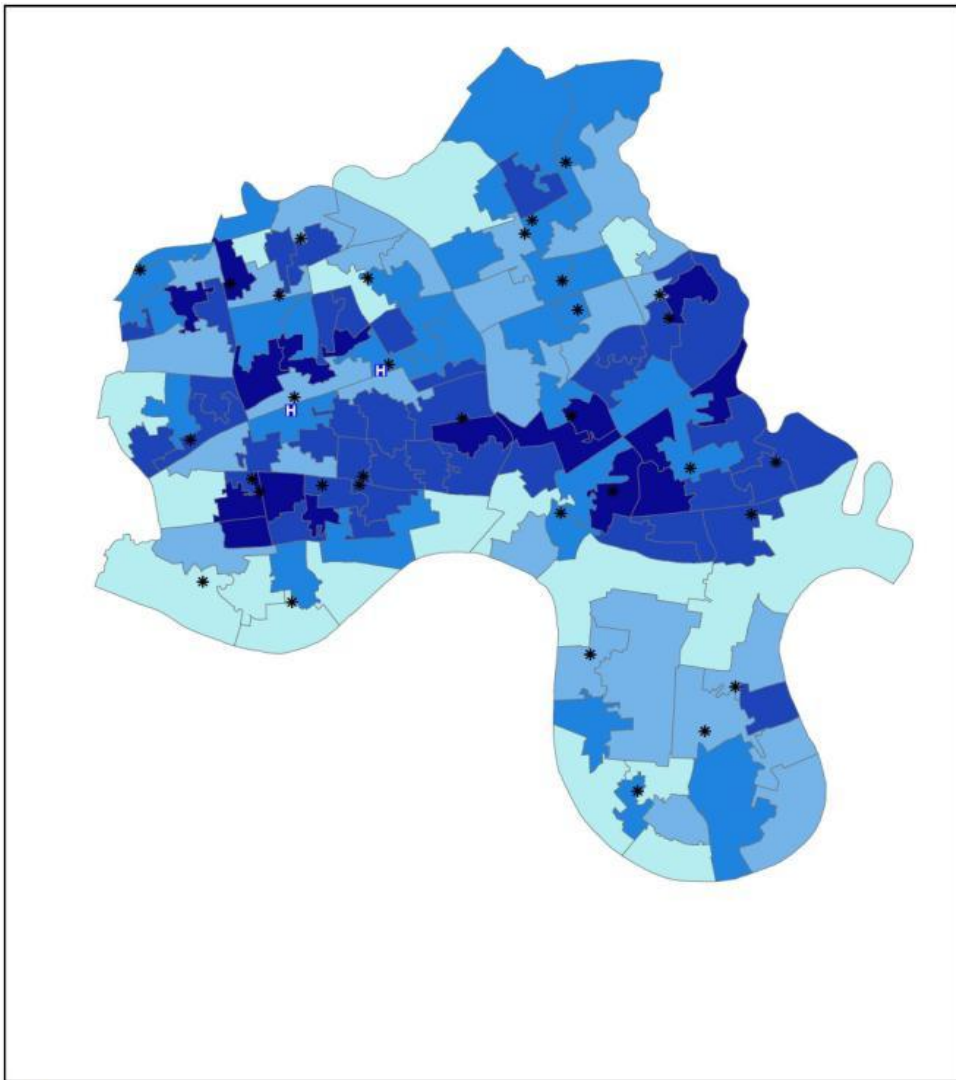**Map 3: Percentage at high risk of type 2 diabetes (heat map)**

Visualising the data as in Map 3 depicts – somewhat more dramatically – a high-risk 'hot' band running west to east through the deprived housing estates and much lower-risk 'cool' areas in the more affluent riverside in the south and park-side in the north of the borough. The heat map is free from the visual lower super output area administrative boundaries that are commonly used in maps of the *basic* type. The resulting map is likely more intuitive for users to interpret due to the colour scheme and there are no boundaries to disrupt the visualisation of diabetes risk.

A second *heat* Map 4 shows the same technique but only for the South Asian at high risk population. It had been hoped that this would illustrate the finding that the South Asian population were at higher risk overall in most areas. However, the technique used to create the heat maps did not allow for the same colours to be applied to the same values between Map 3 and Map 4, and so this comparison could not be made visually. The result was the finding that the prevalence of high type 2 diabetes risk for the South Asian population was concentrated more in the West of the borough. It should be noted that the scale on Map 4 is 9.5-23.7% as compared to 5.1 to 13.8% on Map 3 of the population overall.
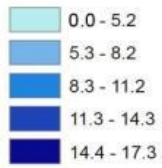
Percentage South Asian at high risk

23.7

9.5

**Map 4: Percentage South Asians at high risk of type 2 diabetes (heat map)**

Map 5 and Map 6 show the prevalence of high type 2 diabetes risk overall and in the South Asian population only across lower super output areas using the *basic* map technique. This method allowed the same colours to be allocated to the same values, with the exception of the fifth quintile which differs in the South Asian population as prevalence extends to 30.3%. These two maps, unlike the two heat maps, made the finding of higher risk in the South Asian population visually compelling.

**Legend:**

- ⊞ Hospitals
- ✳ GP practices

**Percentage at high risk**

- 0.0 - 5.2
- 5.3 - 8.2
- 8.3 - 11.2
- 11.3 - 14.3
- 14.4 - 17.3

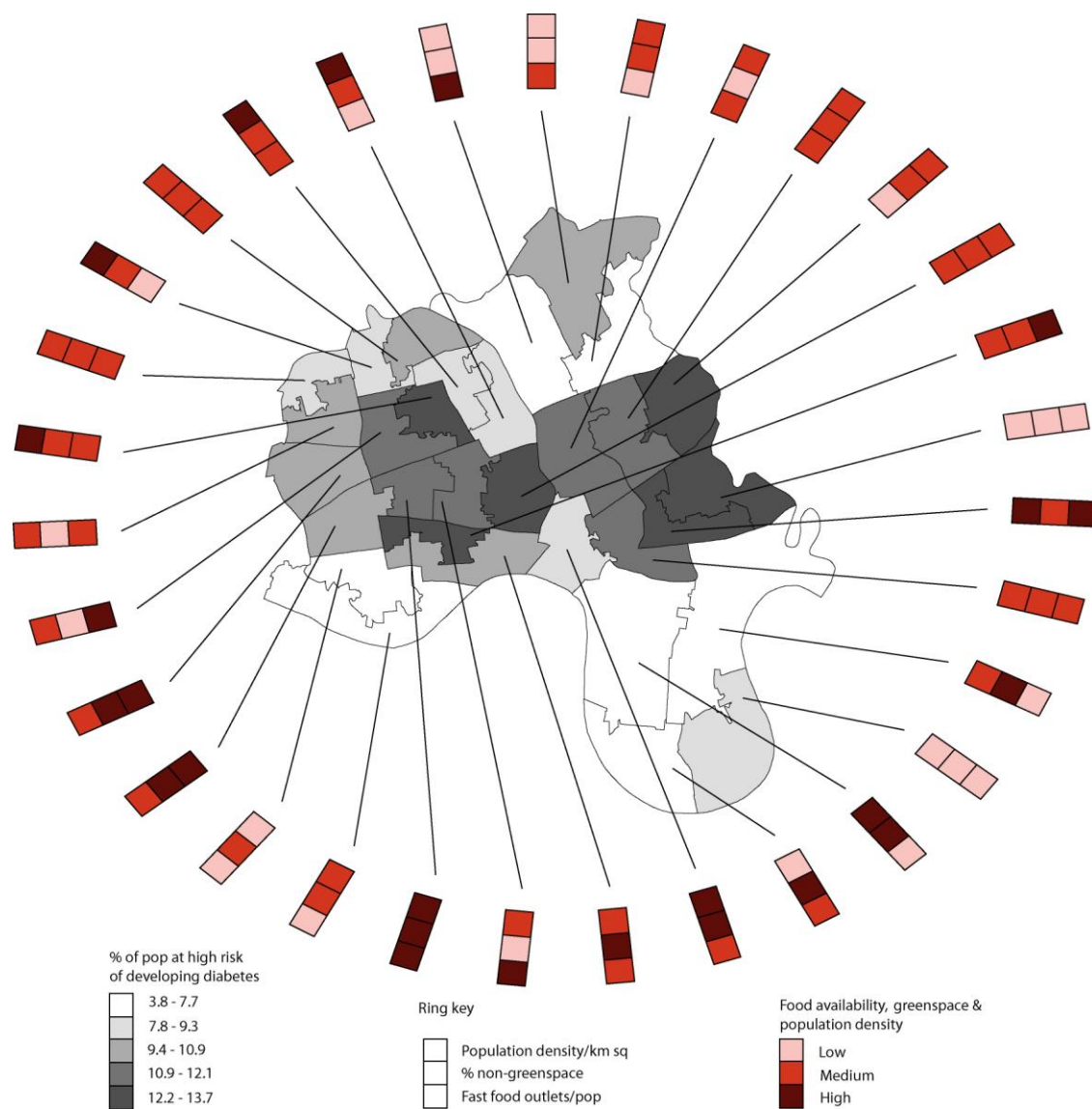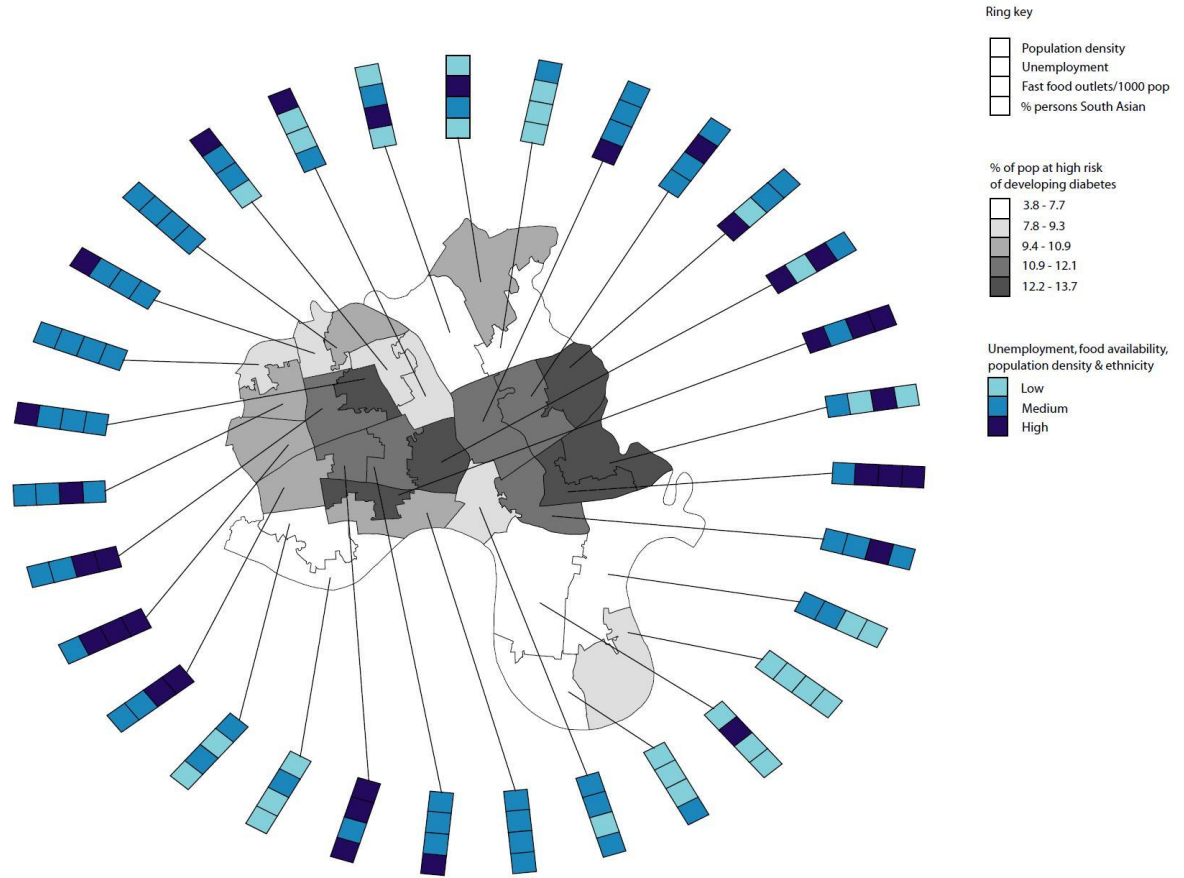**Map 5: Percentage at high risk of type 2 diabetes by lower super output area (2)**
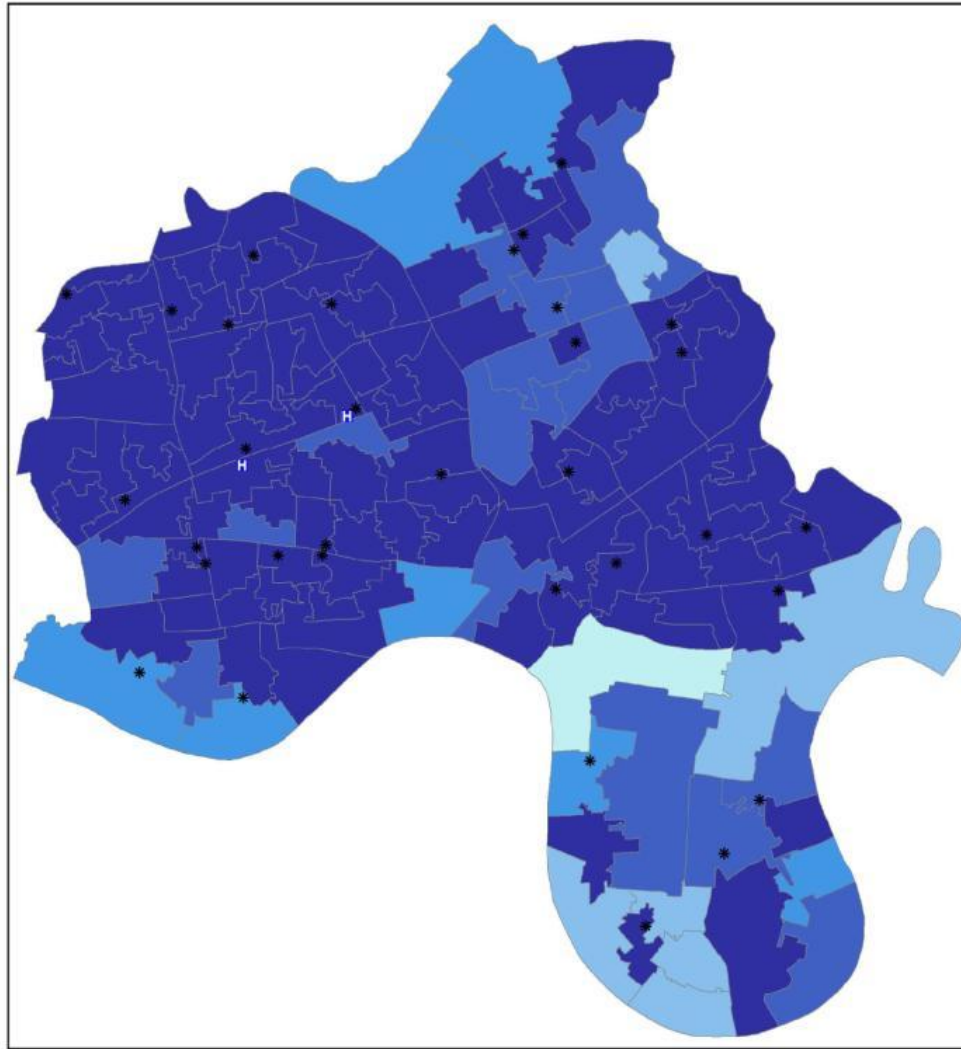
**Map 6: Percentage at high risk of type 2 diabetes by lower super output area (South Asian population only)**

The *ring* Map 7 shows prevalence of high diabetes risk by middle super output area. In this depiction of the data, prevalence of diabetes risk ranges from 3.8 to 13.7%. Each middle super output area is shown linked to a band of three social and environmental indicators which are often suggested to influence poorer health.[141] These are (from the inside out) fast food outlets per head of population, percentage of non-green space and population density per square kilometre. A second *ring* Map 8 was created without %non-greenspace, and added South Asian ethnicity and unemployment (derived from the IMD). This had a similar visual effect, but was not regarded as robust as Map 7 because of colinearity linked to ethnicity and unemployment (which is measured within IMD and Townsend score which is within the QDScore).

Overall, the *ring* maps provided a striking visual display of type 2 diabetes risk in the areas which corresponded to known deprivation and the ring provided a relatively new way of displaying social and environmental determinants of health at a small area level. The ring provides a dashboard of indicators of wider determinants of health that appeared most useful when locally applied to specific population groups of 5,000-7,200 persons. It demonstrates the sort of putative environmental determinants that public health specialists may want to map as part of routine health needs assessment to inform interventions at small area level.

% of pop at high risk
of developing diabetes

3.8 - 7.7
7.8 - 9.3
9.4 - 10.9
10.9 - 12.1
12.2 - 13.7

Ring key

Population density/km sq
% non-greenspace
Fast food outlets/pop

Food availability, greenspace &
population density

Low
Medium
High

**Map 7: Ring map highlighting links to selected determinants of health**

**Ring key**

- Population density
- Unemployment
- Fast food outlets/1000 pop
- % persons South Asian

**% of pop at high risk of developing diabetes**

- 3.8 - 7.7
- 7.8 - 9.3
- 9.4 - 10.9
- 10.9 - 12.1
- 12.2 - 13.7

**Unemployment, food availability, population density & ethnicity**

- Low
- Medium
- High

**Map 8: Ring map highlighting links to selected determinants of health (2)**

Map 9 shows the proportion of people at the level of an individual general practice at high risk of developing type 2 diabetes within ten years. The prevalence of high diabetes risk varied from 4.1 - 16.7% across all three boroughs from the cross-sectional study. This revealed a band of risk, which could be called the 'East London Diabetes Belt'.

This belt of risk stretches from Tower Hamlets in the west, with a high Bangladeshi population, to north-east Newham, where there is a high percentage of South Asian and Black African ethnic groups. Affluent riverside properties in the South and parkside residences in the north show low levels of type 2 diabetes risk.

The 'East London Diabetes Belt' is similar to the 'Diabetes belt' in the Southern States of America.[127] This has high potential to inform the work of commissioners, with a view to taking action to reduce incidence of type 2 diabetes through locality-based interventions. In some areas, almost 1 in 6 adults fell into the 'high risk' category.

Percentage at high risk by general practice

High : 16.7

Low : 4.1
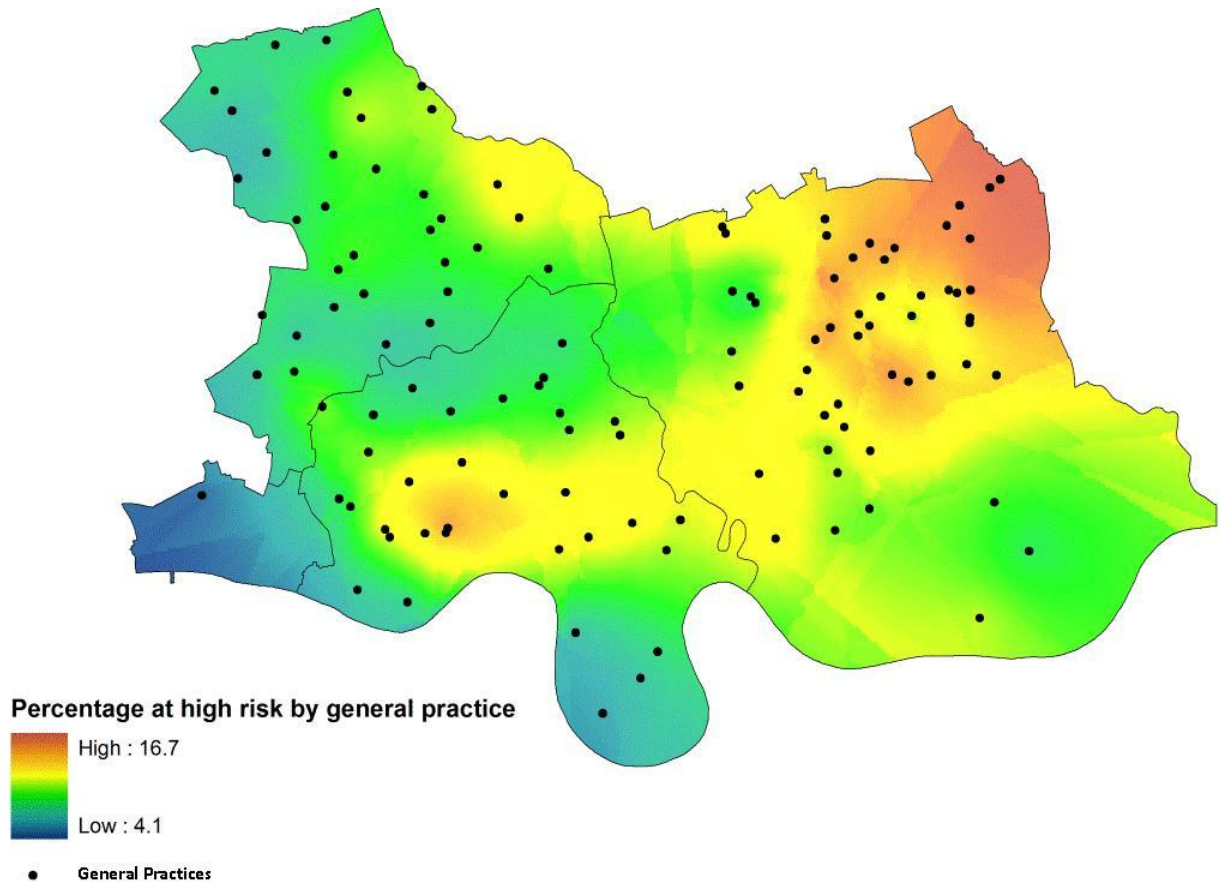
• General Practices

**Map 9: Percentage of patients at high risk of type 2 diabetes by general practice – The 'East London Diabetes Belt' (heat map)**

Map 10 shows Charles Booth's historical map of poverty in London from the late 1800s, which was created using subjective judgements of poverty based on direct observation.
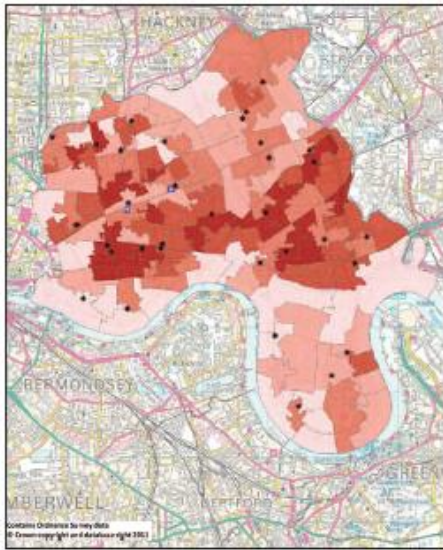
**Classification of poverty**

- BLACK: Lowest class. Vicious, semi-criminal.
- DARK BLUE: Very poor, casual. Chronic want.
- LIGHT BLUE: Poor. 18s. to 21s. a week for a moderate family
- PURPLE: Mixed. Some comfortable others poor
- PINK: Fairly comfortable. Good ordinary earnings.
- PURPLE: Mixed. Some comfortable others poor
- PINK: Fairly comfortable. Good ordinary earnings.

**Map 10: Charles Booth's map of historical poverty in London in the late 1800s**

Map 11 shows a sequence of four maps. First, Map 1 of high risk of diabetes and Map 2 of IMD showing the overlap between risk and deprivation. Second, Map 1 is compared to Map 10. A simple circling technique reveals that there are similarities between areas of poverty in the late 1800s and areas of high risk of diabetes (proxy for deprivation) in 2011.

137

Percentage at high risk of diabetes
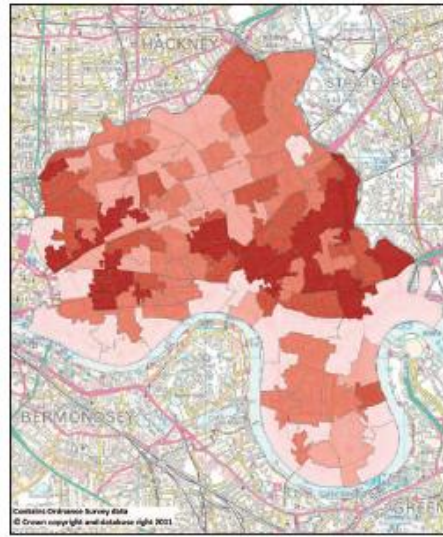
Index of Multiple Deprivation 2010

Classification of poverty

BLACK: Lowest class. Vicious, semi-criminal.
DARK BLUE: Very poor, casual. Chronic want.
LIGHT BLUE: Poor. 18s. to 21s. a week for a moderate family
PURPLE: Mixed. Some comfortable others poor
PINK: Fairly comfortable. Good ordinary earnings.
PURPLE: Mixed. Some comfortable others poor
PINK: Fairly comfortable. Good ordinary earnings.
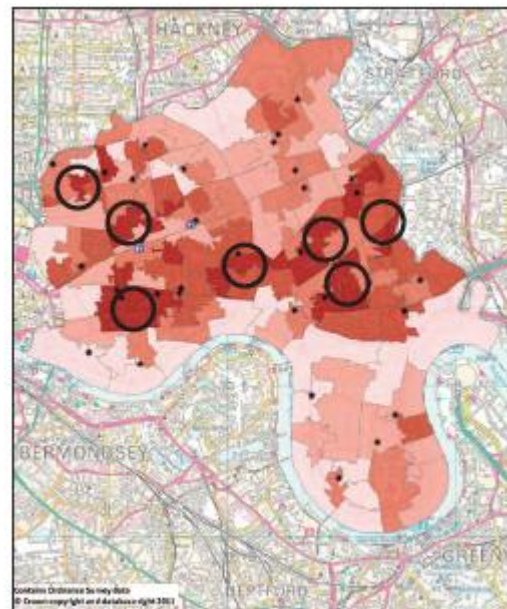
Percentage at high risk of diabetes

**Map 11: Comparison with historical maps of poverty**

This final analysis highlights an important message: despite changing social determinants of health and vastly different relative levels of poverty, a few geographical hotspots of particular need appear to have remained constant.

As was anticipated, the information governance challenges were substantial and were as time consuming as the technical ones. In order to access the data (including postcodes) from general practice records, permission had to be obtained from both the local information governance committee of the PCT and the NIGB. In addition, because this project had a research element, it also required advice from the local NHS Research Ethics Committee and from the university's Research and Development Office (who both deemed the project 'audit'). Potentially identifiable data from patient records had to be handled securely under a protocol advised by the NIGB. This kept postcode information separate from clinical variables with pseudonymised conversion to lower super output area.

Information governance issues were therefore time-consuming and required specialised knowledge and formal permissions, but they were not insurmountable. Furthermore, the process of establishing a procedure for the current project built a stock of knowledge and a network of contacts which would make any subsequent set of permissions and procedures substantially easier to set up.

The practical challenges of undertaking this work were relatively minor. However, this was probably due to a near-optimal local infrastructure. Unusually, there was access to a single

electronic database covering an entire PCT area, due to unique data sharing arrangements between the local general practices, the PCT and the university. Furthermore, the quality and completeness of general practice electronic data across the borough was high. Those seeking to replicate this approach in other parts of the country may need to undertake groundwork to establish a mechanism for data extraction from multiple different computer systems, underpinned by relationships and permission for governance, data sharing and data quality.

Technical challenges included downloading and cleaning the data, which had to be done in several stages due to the size of the files and handling of multiple variables. Conversion of postcode to lower super output area with look-up tables and secure data pairing protocols between datasets 1 and 2 was time consuming. Specialist software was expensive and different versions used between CEG and the Geography Department was inconvenient and resulted in time spent converting files and reducing lines of data, with older software unable to hold as much data. EMIS Web does not keep records of searches performed once an update is installed (which occurs every 4-6 months), so there is a limited time window for cross-sectional analysis.

All geographical work was carried out on a 256-bit NHS encrypted memory stick in the Geography laboratory so that files with lines of patient information were never used outside the CEG except on secure memory sticks. This was time consuming and prevented regular backup of data, which had to be done between two encrypted memory sticks periodically.

140

The technical process of mapping was relatively straightforward - once the data had been prepared, received and decisions made about what maps to create - as I had the expertise of DS to use GIS and Adobe Illustrator. It is unlikely that without these skills high quality maps could have been produced.

**Discussion**

In this study, it was possible to: [a] obtain a near-complete set of de-identified data drawn from an entire borough's electronic primary care records in an ethnically and socio-economically diverse inner city district, [b] use a computer algorithm to determine ten year risk of type 2 diabetes for individuals on this dataset, and [c] use geospatial mapping to highlight dramatic variation in diabetes risk by small area geography and show how social and environmental determinants of health can be effectively displayed and communicated. Information governance and technical issues were challenging but surmountable. The technique of geospatial mapping, as explored through three different formats, may help to meet the rapidly growing need for local health intelligence by planners and commissioners of health services.

Taking a geospatial view of health information such as population at risk of disease complements a traditional statistical approach to such data. Epidemiologists use statistical tests, arithmetic adjustments, and critique causality claims and data. By contrast cartographers use geospatial visualisation, utilise classing breaks (e.g. quintiles), and critique symbolisation.[122] These different paradigms have an important complementary role. Quantitative analysis identifies statistically significant trends; cartography brings meaning and local relevance. Yet merely converting routine epidemiological data into maps runs the risk of oversimplifying complex data and misunderstands the purpose of geovisualisation, which is to represent data spatially. Grouping and classing data for

142

mapping is an interpretive process, and 'points of interest' to which the eye is drawn on a map may or may not correspond to statistically significant relationships between variables as determined by traditional epidemiological approaches.

The key aim in health mapping is not to identify statistically significant relationships, but to gain firstly insight, then understanding, of the ways in which health status varies over space, and to reveal the potential drivers behind this variation.  In this research, by identifying areas of highest prevalence of high diabetes risk by small geographical areas, local general practitioners, public health specialists and planners can be aware of increased risk and possible causes in their locality, so as to target individual and population interventions.  Such 'local' information may be unlikely to emerge from statistical analyses alone.

Individual health is also linked to non-spatial social determinants, and a map of local-level data is most valuable when interpreted in the wider social context. Relative income inequality within the UK is likely to influence weight (and therefore diabetes) via complex pathways.[142]  One example is the 'obesogenic environment' model which encompasses local and national, physical and social environments.[143]  The maps presented here are ideally considered with this context in mind.

Resources and skills in handling health information in order to commission new interventions and services may be limited, particularly where they relate to dual

responsibility of both local authorities and health providers for the health of local populations. Geospatial mapping offers one option to address these deficiencies and present diverse information about health and its wider determinants in an accessible format to support commissioning and planning expertise. It is possible, though somewhat speculative at this stage, that investment in the skill base needed for this approach may be cost effective in the longer term.

The mapping study is probably one of the first to use routinely collected, local individual electronic patient data to generate high-quality small-area maps of disease risk across an entire borough. A significant strength of the study was the quality and completeness of the dataset from which the geospatial maps were derived. There was up-to-date data on the majority of the target cohort (aged 25-79 years) across the whole of Tower Hamlets.

The completeness of data capture in the study was attributable to a number of things (similar to that for the cross-sectional study except that in Tower Hamlets all these relationships were generally considered to be stronger and more fully developed): [a] existing partnerships between the university and the NHS, [b] a 20 year history of using electronic medical records in local general practices, with standard data entry templates for performance monitoring, audit and needs assessment, [c] existence of local data sharing agreements and information governance infrastructure for overseeing the use of electronic personal medical data, and [d] the fact that 35 out of 36 general practices in the borough

used the same computer system (EMIS) which was compatible with the chosen diabetes risk algorithm (QDScore), and 33 out of 35 shared postcode.

A potential limitation of the study is this uniqueness of the local context. In order for the method used here to be successfully reproduced by others, a number of conditions need to be met. First, effective data sharing agreements must be in place and a high degree of trust is necessary between all parties. Second, the general practice records of a whole population need to be accessible and the quality of relevant data fields on these records (completeness, accuracy and consistency of coding) must be high. Third, the method requires that patients registered at a particular general practice live in the same district. This was not the case for 1,813 (1.1%) individuals in this study. In some other localities this discrepancy might be far greater. Fourth, the task of downloading and cleaning data and geographically mapping disease risk required an advanced set of skills and took many hours of input from a data analyst (RM), public health specialist (DN) and human geographer (DS). It is some way off for a set-up whereby planners or general practitioners can simply hit the 'map it' button on their consoles to produce maps like the ones illustrated in this thesis.

Two areas in this study produced much further discussion with colleagues and I regard them as special interest topics: [a] information governance issues associated with geospatial mapping, and [b] how the media interpreted the findings from the maps.

**Information governance and geospatial mapping**

Changes in the technical infrastructure of general practitioner electronic patient records create the potential for analysis of previously unavailable individual health data for research and audit. Such 'secondary uses' of data collected largely or wholly for the purpose of individual patient care raise substantial technical, security, ethical and civil liberties issues. Concern about protecting patient data is central to information governance in the NHS. With the increasing non-standardised use of mapping for data analysis and presentation to commissioners, new information governance challenges are emerging.

During the course of preparing the geospatial maps advice supplied by the NIGB was followed on how to handle health data for small numbers of people, so as to protect against any person being identified. However, there were specific issues about *how* to conceal small numbers of persons in one of the maps. This was resolved after a meeting with the NIGB, but it revealed that specific published guidance on protecting people from being identified through geospatial mapping were lacking, and many rules of thumb were being used. I decided with my colleagues to write a detailed academic paper with the NIGB on these information governance issues and include the techniques that could be used to protect against identifying people in geospatial maps, and develop an assessment of risk framework for researchers to use. This academic paper has been submitted to the journal BMC Public Health and is included as Appendix 7. Although I conceptualised the idea for this paper, brought the authors together and produced the first draft, DS, the geographer I

had been working with, took the lead on developing the manuscript and themes contained within. Below is a description of some of the early issues that led on to the much more detailed piece of work with a group of experts in the academic paper. I hope the paper may eventually be used as the basis for a government endorsed guideline for researchers.

When data is aggregated to small areas (often lower super output areas in England), there may still exist a potential breach of data protection if the number of individuals within an area are below a threshold number. Techniques to handle this problem can be applied from confidentiality guidance for small area data and health statistics which despite being primarily intended for statistical tables have the potential to be applied to geospatial maps.[144][145]

Firstly, it is important to determine a threshold number of subjects likely to be identifiable in any one area based on the condition being studied (e.g. 3, 5 or 10) and use a denominator as large as possible for creating the map. Once map creation has taken place several possibilities exist for suppressing cells that fall below the threshold: [a] leaving cells blank (care needs to be taken to ensure that the number cannot be calculated from other cells by differencing, although this is less likely to be significant in maps where quintiles and ranges are more commonly used), [b] amalgamation of adjacent cells such that numbers rise above the pre-determined threshold, [c] rounding small numbers up to the threshold or moving up to the next geographical level (although this may significantly change the resulting map), and [d] using ranges of numbers or percentages, for example, converting 3 to 0-10.

Map specific problems need to be borne in mind. For example, suppressing a geographical cell by colouring it white or marking it with an X is more complex than the equivalent manoeuvre in a data table, since it draws attention to the fact that small numbers are present in that geographical location.

Other guiding principles could include: [a] an assessment of the sensitivity of the data, [b] who will ultimately have access to the maps, [c] whether individual people could reasonably be identified, and [d] searching for other maps of the same data and considering whether small numbers could be calculated by differencing or a similar technique. Additionally, even if it is intended that maps will only be available amongst health service planners, the potential of maps to become publicly available once they have been transferred electronically between different parties is high. Because of their highly visual nature compared to data tables, extra caution in electronic transfer of maps should be applied.

Information governance issues in mapping become especially sensitive when the unit of analysis is small, the number of affected individuals is small (making it a real possibility that someone could be identified), and the condition is sensitive or stigmatising (e.g. teenage pregnancy or mental health). Possible solutions exist but there are no agreed standards.

The necessity of geovisualisation for data display is growing with the shifting structure of the NHS, and a standardised and regulated approach to creating maps of health data would allow for greater consistency in outputs, aiding interpretation between user groups and suborganisations, as well as protecting patient confidentiality. In the absence of formal guidelines for governance of mapping, there is a risk that case law (which is likely to accumulate via atypical and highly sensitive cases) will determine the use of geospatial maps for health planning purposes. The research community should urgently seek to rectify this potential gap of comprehensive published guidance on how to handle patient level data for geospatial maps.

A more detailed and refined discussion of these issues and further themes relevant to this area can be found in Appendix 7.

**How the media interpreted the findings from the maps**

Having worked at the Department of Health with the Chief Medical Officer I had acquired some experience previously of working with the media, including: [a] how to prepare a press release, [b] thinking through how the media would interpret reports, research and press releases, and [c] preparing lines to take prior to interviews to ensure that the interviewee stays on message and is not deflected into making a mistake or falling into the trap of inadvertently providing journalists with a sensational (and often inaccurate) headline.

149

I had not initially thought about how this experience would play out during my academic attachment at QMUL. Following the publication of the systematic review, Professor Greenhalgh suggested preparing a press release. This was slightly delayed as the BMJ published the paper without telling us in advance and so the press release was issued on the day of publication, which attracted little interest as journalists prefer to run the story the same day as the paper is released.

This had, however, whetted my appetite for applying previously learned media skills in an academic context, and I was keen to try and do this with the geospatial study. I thought it might make a news story if we were to publish the BMJ Open paper of the geospatial maps and the special report on diabetes risk from CEG on the same day. I suspected this would be difficult to coordinate, yet wondered whether it might be of media interest as there would be a combination of a 'British Medical Journal' publication and an in-depth report from the 'East end of London' revealing a very high level of risk of diabetes.

I met with the press officer (Kerry Noble) with whom I had worked to produce the press release on the systematic review. She agreed that the story was of interest, and appreciated my desire to highlight risk of diabetes in the East end to help with preventive efforts, but felt it was lacking an interesting enough angle to achieve pick-up from the major national news providers.

I needed a fresh approach to the research findings. As I had been walking to work at the university that day to meet the press officer, I had an idea which eventually resulted in a flurry of national media interest in maps of diabetes risk in the East end of London.

I live in Bow, in Tower Hamlets (in a lower super output area with a risk of type 2 diabetes of 0.0-5.4%). Most days I walk down the well known A11 for about 2 miles to work at the university or PCT. This road is better known by its formal names, being called Bow Road, becoming Mile End Road, then Whitechapel Road, and finally Aldgate High Street. There was apparently a suggestion to rename it the Olympic Boulevard as it is the main route from The City of London to the London 2012 Olympic Site. Each time I walk down that road I am reminded of the deep inequalities and health problems inherent to life in the East end. The North side of Bow Road and Mile End Road are well known for affluent terraced housing epitomised by Tredegar Square. The South side has highly deprived high rises and housing estates. From Mile End Road and onto the early part of Whitechapel Road both sides of the road are awash with small fast food takeaways (mostly deep fried chicken), although there is one Subway, and two Nandos, but no other chain restaurants. On this part of the road both sides contain highly deprived housing, with the occasional pocket of private housing. Many people (sometimes I think most) are smoking. Often I will walk past a group of people on the street drinking super strength lager out of cans, and others who appear to be homeless.

Visually the ethnic diversity is striking.  For example, in addition to the diverse daily market on Whitechapel Road, I sometimes overhear four or five different languages being spoken on one journey.  Mosques lie tucked away behind unremarkable shop-like entrances, and there is often a loud Christian street preacher outside Mile End or Whitechapel tube stations.

At the start of Whitechapel Road there are two statues in quick succession in honour of the famous reformer, missionary, and founder of the Salvation Army, William Booth.   As I walked past these two statues on a cold day in January 2012 on my way to meet the press officer, I was reminded of one of the other famous 'Booths', Charles Booth, who had created well-known maps of poverty in London in the late 1800s.  This gave me an idea.  What if there was a similarity between the areas of diabetes risk and poverty that we had revealed and the historical maps of poverty from the late 1800s.  I thought this could well be a hook that the media would latch onto and provide the platform to discuss risk of diabetes today.

With the help of the press office we obtained the Booth maps from the London School of Economics[139] and there was indeed visual similarity between areas of historical poverty and modern day deprivation and risk of diabetes.  The comparison was slightly tenuous. The scales of the maps were not exactly the same, the Docklands had far fewer residential properties in Victorian times, and the methods used were very different (observation versus highly sophisticated electronic spatial analysis).  Yet, despite these limitations a pattern was

evident and after discussion with colleagues I decided to use the Booth map comparison to highlight modern day risk of diabetes. The press officer was in favour, and drafted a press release[146] which I edited. It is shown in Box 2 below:

---

***Modern health mapping shows how poverty and ill health persist over 100 years***

*Researchers from Queen Mary, University of London are aiming to improve the health of Londoners by combining a century-old mapping technique with up-to-the-minute technology.*

*Using type 2 diabetes as their example, the researchers have compiled detailed maps of east London highlighting the geographical 'hotspots' of disease risk.*

*The maps, which are published today in BMJ Open reveal startling similarities to the renowned 'poverty maps' created in the late 19th Century by Victorian reformist, Charles Booth.*

*The researchers chose to study type 2 diabetes risk because it has well-known risk factors and is preventable. It is strongly associated with poverty and South Asian ethnicity, both of which are common in east London today.*

*The aim of the project is to help local authority and NHS services to tackle poor health by directing efforts where they are most needed. Although the study examined the London boroughs of Tower Hamlets, Hackney and Newham, the researchers say that the same technique could be applied anywhere in the country, and to other diseases.*

*Unlike the Booth maps which were based on observation, the new study uses an entire set of electronic records from GP surgeries in the area. This very precise information means that the maps are much more accurate and will be useful to individual GP surgeries.*

*Electronic records from over half a million people were included in the research. Each was assessed for risk of developing diabetes using a well-established prediction tool, the QDScore.*

*People were categorised as 'high-risk' if they were found to have a one in five or greater risk of developing diabetes within ten years.*

*Overall around ten per cent of the adult population fell into the high-risk category. However the maps showed 'hotspots' where up to 17 per cent were at high risk. Further analysis showed that these hotspots were associated with areas of poverty.*

---

*These hotspots were surprisingly similar to areas of poverty highlighted in Booth's maps from over 100 years ago.*

*The study was led by Douglas Noble, a Public Health Doctor and Lecturer at Barts and The London Medical School, Queen Mary, University of London, and published in BMJ Open with additional material in a full report aimed at the NHS and Public Health specialists.*

*Dr Noble said: "It was no surprise to see that diabetes risk is high in areas where poverty was high. What was surprising was that some of these pockets of deprivation and ill-health have persisted for over 100 years.*

*"But unlike in Booth's time, we now know how diseases like diabetes can be prevented. Using electronic records to create maps like these throughout the country could improve health and save money for the NHS.*

*"When you think of what life was like in the East End in the late 1800s it's extraordinary what the NHS and public health professionals have achieved, often with limited resources. But there's more still to do, and we hope this detailed information will help to reduce risk of diseases like diabetes"*

*The research also looked at known risk factors and could show where a lack of green space or a proliferation of fast food outlets could be contributing to ill-health.*

*Trisha Greenhalgh, Professor of Primary Health Care at Queen Mary, University of London, also worked on the report. She said: "Health mapping has enormous potential for the NHS, especially with a disease like type 2 diabetes which we know can be prevented by keeping a healthy weight and staying active.*

*"This study, which concentrates on three of the 'Olympic boroughs', highlights the dire need for a major and lasting Olympic legacy to improve health and longevity in east London."*

*Steven Cummins Professor of Urban Health at Queen Mary's School of Geography commented: "Population health has vastly improved over the last 100 years. However, as these maps starkly illustrate, a century of social, economic and physical change has failed to eliminate underlying geographical inequalities in disease in east London."*

*This work was funded by Tower Hamlets, Newham, and City and Hackney primary care trusts and by the National Institute for Health Research. The National Information Governance Board advised on data protection issues.*

**Box 2: QMUL press release on diabetes risk Feb 2012**

I thought the release read well, although I had a slight reservation about the headline being solely about the historical aspect, but nevertheless I decided to go with the experience of the press office, on the basis that all the main information about diabetes was included within the release. The BBC was the first major organisation to carry the story, shown in Box 3 below:



**Box 3: Coverage of diabetes risk research by bbc.co.uk**

*Reproduced from http://www.bbc.co.uk/news/uk-england-london-17062735*

The BBC led with the historical angle as the headline. They included a pdf of five of the maps so that readers could make a comparison of the findings themselves. The report was accurate and I was pleased that they covered both ethnicity and deprivation as major risk factors in the opening sentences, as well as the aim of informing commissioning. The latter part of the article included a black and white photograph of poor looking East Londoners in the 19th century.

The next major press agency to cover the story was the Daily Mail Online[147], shown in Box 4 below:



**Box 4: Coverage of diabetes risk research by Daily Mail Online**

*Reproduced from http://www.dailymail.co.uk/news/article-2102485/Updated-maps-Londons-poorest-areas-epidemic-junk-food-diabetes-streets-Victorians-died-malnutrition.html*

This newspaper took a more sensational angle consistent with the approach of the tabloid press.  At first I was somewhat dismayed by the headline as it initially struck me as inaccurate, but after careful thought I realised that in fact it was quite a clever interpretation of the research findings.  Unlike the BBC the headline did not focus solely on the historical angle, but included information on the determinants of health.  I assume the journalist had studied the *ring* maps and noticed that in some of the areas of high diabetes risk there was also a preponderance of fast food restaurants.  The headline highlights the geographical comparison between risk factors for poverty today (fast food leading to diabetes), with risk factors for poverty in the Victorian Era (malnutrition leading to death).  Both themes are encapsulated with the prefix: *'The changing face of poverty'*.  Drawing attention to the paradox of lack of food and poverty historically, with too much of the wrong type of food and modern-day poverty was clever, eye-catching, and on-message with mainstream public health promotion.

The article itself concentrated on this angle of the changing nature of the social determinants of health, and included an interesting box on the life of Charles Booth.  Several pictures were chosen, including: two of the maps, a picture of a poor looking Victorian family, a Victorian street scene with poor looking children, a picture of Charles Booth, and a picture of modern day young people consuming fast food.  The last picture which was put alongside the scene of poor Victorian children is shown below in Box 5.

Hard times: While many suffered ill-health due to starvation in Victorian Britain (like these children in 1860) modern youngsters are at risk of developing diabetes through addiction to junk food

**Box 5: Picture from Daily Mail Online**

*Reproduced from http://www.dailymail.co.uk/news/article-2102485/Updated-maps-Londons-poorest-areas-epidemic-junk-food-diabetes-streets-Victorians-died-malnutrition.html*

Whilst the Daily Mail did mention other risk factors, the focus was mainly on fast food, epitomised by the right hand picture in Box 5. It highlighted the damaging effects of a reversible risk factor (fast food consumption) to health through the link to development of diabetes. I was pleased that a reversible risk factor was chosen by the newspaper, as it shows how individual choices can affect risk of diabetes. It was also noticeable that the three youths pictured in Box 5 are of a lower risk ethnic group, and are not visually clinically obese. This sends out a satisfactory public health promotion message i.e. you

don't need to be already overweight or of a certain ethnic group to be at increased risk of diabetes - poor diet is a risk factor in and of itself because of the potential it has to cause obesity in the long term.

A local newspaper called East End Life[148] also featured the research in print (and available online), as did the Daily Telegraph in their print edition only. BBC 1 London television covered the story on the evening regional news, and a journalist from The Economist interviewed me in connection with a larger piece on poverty in London to be published later this year. I turned down a radio interview with BBC London *Drivetime* as I had concerns after speaking to the journalist on the telephone that they were going to focus on immigrants living in poverty, telling me they'd done: 'a lot on diabetes recently'. I felt that was an unhelpful angle. Table 13 shows an analysis of the emphasis placed on various themes from the original press release from the news agencies that covered the research online.

| Major themes | QMUL Press Release | bbc.co.uk | Daily Mail Online | East End Life |
|---|---|---|---|---|
| **Historical poverty** | +++ | +++ | +++ | +++ |
| **Risk of developing diabetes** | +++ | +++ | +++ | +++ |
| **State of the art mapping technique** | +++ | + | + | ++ |
| **Diabetes risk factors** | +++ | +++ | +++ | +++ |
| **Aiding commissioning of prevention** | +++ | +++ | ++ | ++ |
| **Using research methods in other places** | +++ | + | + | + |
| **Diabetes risk scoring** | +++ | | ++ | ++ |
| **Primary care electronic records** | +++ | + | + | + |
| **The Olympic boroughs** | +++ | +++ | | |
| **Geographical inequalities** | +++ | | | |
| **Fast food** | + | | +++ | |

**Table 13: Analysis of media coverage of diabetes risk**


Overall the media coverage was very good, and highlighted the major issues of risk of diabetes and common risk factors. The decision to use the comparison with historical maps as an entry point to a wider discussion on diabetes worked well. In Section 5 I discuss further how the media coverage has helped to contribute to further preventive action locally.

**SECTION 5: Conclusions**

**Chapter 9: Summary, implications and personal reflection**

**Key findings**

*Section 1* highlighted how type 2 diabetes is on the rise, with increases in prevalence predicted internationally, nationally, and locally. East London faces vastly increasing rates of diabetes by 2020 with the number of diagnosed cases expected to rise from approximately 40,000 to 80,000.[109] Risk of type 2 diabetes is closely associated with development of established disease and could reasonably be expected to rise by the same factor resulting in 1 in 5 adults at high risk of type 2 diabetes by 2020 in East London. Urgent public health action needs to be taken not just to improve early diagnosis and management of established diabetes, but to reduce risk of developing diabetes across the whole population. In face of this, robust methods for identifying those at risk of diabetes are essential.

The systematic review of diabetes risk scores in *Section 2* considered the performance and impact of risk models and scores for predicting risk of type 2 diabetes in adults without known diabetes. Results showed that there were some diabetes risk models and scores that could be relatively easily applied to routinely collected data, such as that contained within electronic primary care records. However, out of 145 models and scores, performance and

impact was very mixed, with many scores not making any impact in the real world. Much more work is needed to assess impact and usability in the long term.

In *Section 3*, using one of the risk scores from Section 2 (the QDScore) on an entire population's electronic primary care records, I explored how feasible it was to aggregate, describe, and stratify diabetes risk in a way that meaningfully informs locality-based needs assessment and service planning. This was successfully done, although there were various hurdles, including statistical colinearity, complex data handling issues, and a lack of engaging presentation techniques. The results showed that 1 in 10 adults were at high risk of type 2 diabetes and, in particular, there was significant association with cardiovascular morbidity, ethnicity and deprivation.

Finally, in *Section 4*, research was also undertaken on the feasibility of geospatially mapping an entire population's risk of type 2 diabetes in a way that could lead to engagement of commissioners on the usefulness and applicability of the findings. This was also successful, despite information governance challenges, and judging by the media response, has high potential to highlight risk of type 2 diabetes to a wide audience. Further more detailed formal qualitative research of impact would be a logical next step.

**Implications for policy, practice and research**

Geoffrey Rose, the famous public health professor, originally prioritised population interventions over targeting individuals.[149-151] Although both strategies are not in conflict, in response to interventions aimed at individuals of which only a few benefit, Rose concludes: [150]

*'We are therefore driven to consider mass approaches, of which the simplest is the endeavour to lower the whole distribution of the risk variable by some measure in which all participate.'*

As Rose also suggests this may result in *the prevention paradox*[149-151] i.e. benefit conferred by the *mass approach*, may confer little gain for those at high risk.

The Marmot review *Fair Society, Healthy Lives: A Strategic Review of Health Inequalities in England* recommended *proportionate universalism* for tackling health inequalities, defined as follows:[152]

*'Focusing solely on the most disadvantaged will not reduce health inequalities sufficiently. To reduce the steepness of the social gradient in health, actions must be universal, but with a scale and intensity that is proportionate to the level of disadvantage.'*

Applying and drawing on these two approaches to the findings in this thesis, it would appear that the weight of healthcare resources and public health interventions focus mainly on those who have established diabetes and those with undiagnosed disease. Efforts mainly revolve around controlling disease, with monitoring of biochemical parameters such as HbA1c. This approach alone is unable to deal with increasing prevalence. The weight of public health intervention needs to shift as on the population curve shown in Figure 9.
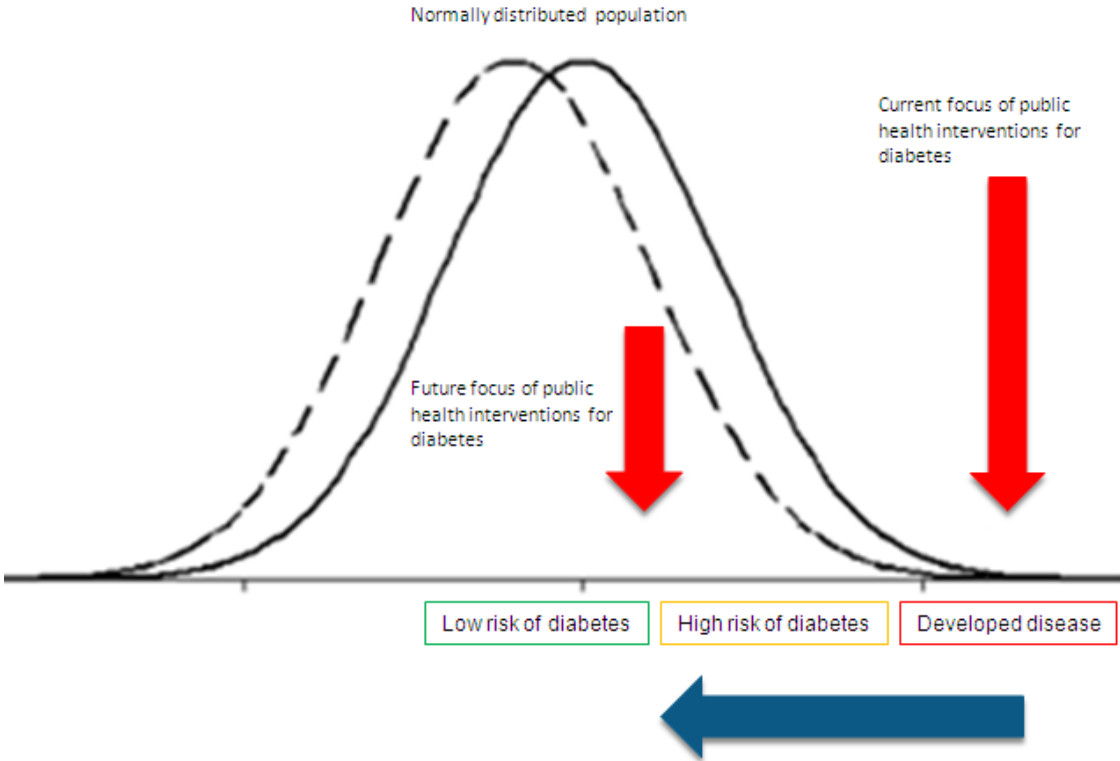


**Figure 9: Moving the focus of public health interventions to prevent diabetes**

By targeting those at risk of developing type 2 diabetes, disease prevention has the potential to reduce incidence, prevalence, morbidity, mortality, healthcare use and costs, and increase

quality of life.  This type of paradigm shift requires high-level health policy to drive whole system reform.

The need to reduce the prevalence of non-communicable diseases has been recognised internationally by the United Nations.  The growing burden of diabetes, alongside other non-communicable diseases such as cardiovascular disease, respiratory disease and cancers, has been met with a call from the United Nations General Assembly at the 2011 United Nations High Level Summit on Non-Communicable Disease.   They have called for a strengthening of national policies and health systems, population wide interventions, primary care services and disease monitoring across the whole population.[153]

Closer to home in the UK, the Foresight Report on obesity (the major risk factor for diabetes) stated: '...*a bold whole system approach is critical – from production and promotion of healthy diets to redesigning the built environment to promote walking, together with wider cultural changes to shift societal values around food and activity.*" [154]

More specifically in May 2011, NICE produced guidance on population and community interventions aimed at preventing diabetes.[13]  Further guidance, currently in draft form, also by NICE, specifically addresses interventions for individuals at high risk.[155]  This offers the possibility of shifting the focus of public health efforts towards prevention both nationally and locally.

Several principles relevant to many public health areas underpin the NICE guidance for populations and communities: [a] behaviour change through education, [b] emotional support and planning, [c] weight management through healthier eating (e.g. 'five a day') and interventions aimed at weight reduction that are measured, specific and individual, [d] physical activity including 30 minutes five days a week, and [e] cultural sensitivity to ensure that interventions take account of language and literacy, educational barriers, religion and cultural norms. Eleven specific recommendations are made. Many are expressed in generic terms, covering strategy, health promotion, education, physical activity, healthy eating and targeted prevention. These are summarised in Table 14.

| National recommendations | National and local recommendations | Local recommendations |
|---|---|---|
| Working towards integration of strategies for all non-communicable diseases, including partnership working, focusing on cross-cutting risk factors, addressing demographic disparities, collating basic epidemiology, improving availability of resources. **Strategy** | Educating workers who have a role in health promotion. **Education** | Identifying high risk communities by utilising routine health intelligence through joint strategic needs assessments, locating existing interventions, and finding community organisations with potential for health promotion functions. **Targeted prevention** |
| Promoting healthier eating habits, including working with the private sector, manufacturers, caterers and retailers to promote healthier foodstuffs. **Healthy eating** | Delivery of culturally sensitive health promotion information to the entire population which tackles misunderstandings and promotes healthy eating and exercise. **Health promotion** | Locally led strategy formation involving best use of evidence and local cost effectiveness knowledge, environmental change, and targeting high risk groups. **Strategy** |
| Promoting exercise, including highlighting recommended daily amounts, changing the built environment, and tracking progress. **Physical activity** | | Specific interventions for high risk community groups involving partnership, outcome measures, education and training of workers from a range of backgrounds, and appointing community champions. **Targeted prevention** |
| | | Adapted and targeted national health promotion messages. **Health promotion** |
| | | Promoting healthier eating habits by working with local partners and businesses, and making information available about entitlement to discounts and menus. **Healthy eating** |
| | | Promoting exercise, including working with local planning departments and leisure services, local employers, and highlighting recommended daily amounts. **Physical activity** |

**Table 14: NICE populations and communities guidance on interventions for preventing diabetes**

NICE estimate that interventions costing £10 (for the whole population) and £100 (for Black and minority ethnic groups) per person, which returned a mean weight loss of 0.25kg and 1kg respectively would be cost effective at a cost per quality-adjusted life year threshold of £20,000. Weight loss interventions in Black and minority ethnic groups need participants to lose 3-4kg to achieve fiscal savings. This led NICE to the conclusion that less expensive population wide interventions have to be combined with effective individual interventions targeted at those at high risk. This approach is consistent with the Marmot review's recommendation of proportionate universalism i.e. in order to tackle health inequalities across all of society, public health action should be appropriate for everyone, but proportionally more for those whose need is highest.[152] At a policy level the findings in this thesis offer an approach that could be used to help achieve this dual approach. As summarised in Chapter 2, Black and minority ethnic groups and people from deprived areas have been shown to be at much greater risk of type 2 diabetes. Electronic record analysis and risk scoring could allow health planners to identify *both* 'locality hot-spots', specific high-risk individuals *and* the level of risk across the entire population, with a view to achieving proportionate universalism.

A second set of NICE guidelines specifically considering individual interventions for individuals at high risk is currently in draft form.[155] The draft recommendations in this report consider two broad areas: [a] identification of high risk people, and [b] individual interventions to reduce risk. Draft recommendations include healthy eating, physical activity, targeted interventions, pharmacological therapy, surgery and risk identification.

As described in this thesis the latter has significant potential to accurately identify high risk individuals, and such a system could be incorporated into the NHS Health Checks programme, which currently targets 40-74 year olds for an assessment of vascular disease.[43] This programme already includes identification of diabetes for some high risk individuals. Using a validated tool such as the QDScore could result in more accurate testing of high risk individuals for further follow-up and interventions.

Figure 10 shows the clinical pathway for NHS Health Checks.[43] It already includes a *diabetes filter*. Replacing the current filter with the QDScore could potentially allow more accurate triage of high-risk individuals for further testing/interventions. I have already started to take this work forward with the local PCT partners, with a view to incorporating QDScore in NHS Health Checks locally from the next financial year.
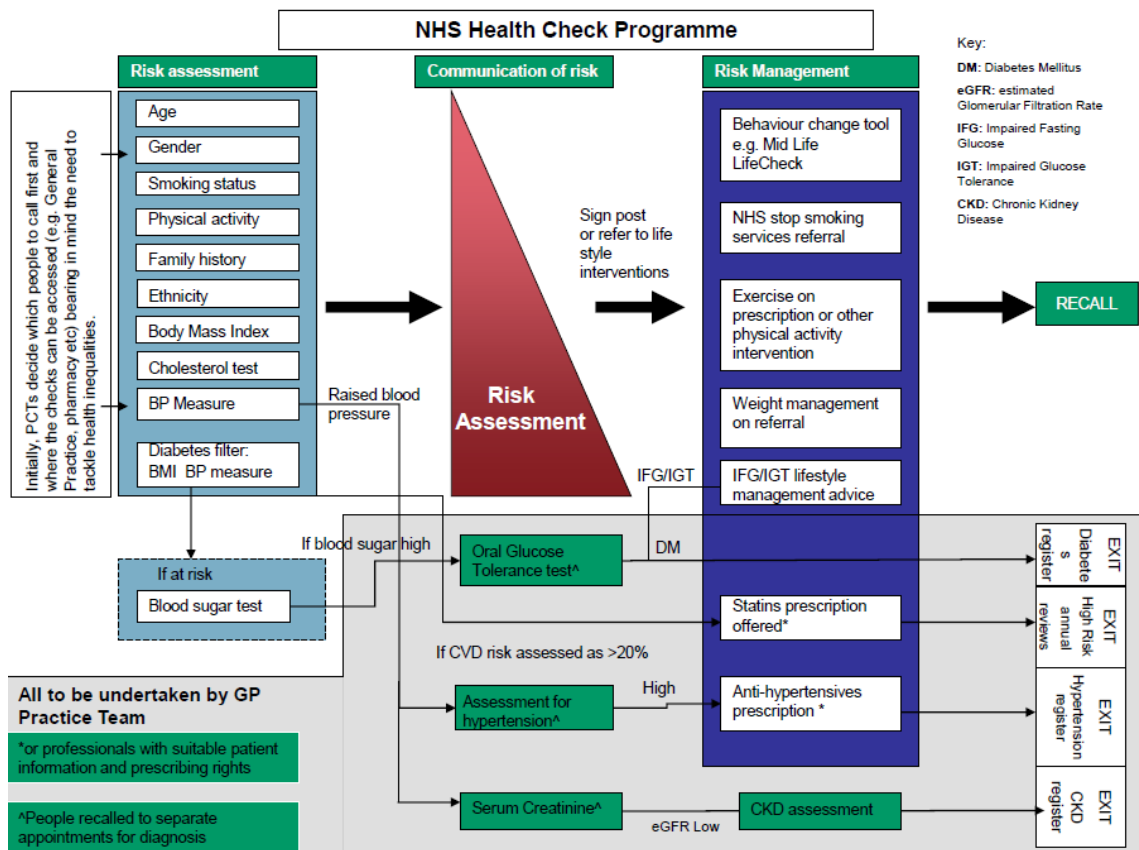
**Figure 10: NHS Health Check Programme**

*Reproduced from*

*http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_098410.pdf*

Essential to using QDScore within NHS Health Checks will be its automatic incorporation within the electronic GP record. This would mean locally that rather than a GP calculating it manually and entering every patient risk factor, EMIS Web would contain the algorithm,

extract the risk factors, and present the risk score directly to the GP. This is currently under development.

On the back of the media coverage, and with the advent of the QDScore, which has just recently been renamed QDiabetes, as a downloadable app on *apple.com*, I have begun to explore using it more widely locally. This has involved early discussion around incorporating it on every bedside monitor in the acute hospitals. I have also been in discussion with a diabetes charity, that aims to raise awareness of diabetes, and that had approached Tower Hamlets PCT. They wanted to perform random point of testing for blood glucose around the borough to raise awareness for diabetes. I met with them and the Director of Public Health, and we have suggested they use the QDScore instead to highlight risk, and aim to park their vans (with testing stations and accompanying health promotion materials) in the areas identified as high risk in the maps in this thesis. This discussion is ongoing.

Delivering diabetes prevention in people who are not suffering from any disease requires skills which traditionally-trained clinicians may not possess.[156] Almost nothing is known about the reach, uptake, practical challenges, acceptability and cost of preventive interventions in high-risk groups in different settings.[27] The relative benefit of detecting and targeting high-risk individuals rather than implementing population-wide diabetes prevention strategies is also largely unknown.[28] These are potential areas of further research to assess most effective prevention strategies and approaches.

The preliminary findings from the impact studies covered in the systematic review also suggest that not everyone at high risk is interested in coming forward for individual preventive input, nor will they necessarily stay the course of such input. Researching which factors buck this trend could result in improved prevention in the future.

We know from cohort studies that early detection of established diabetes improves outcome, though the evidence base for screening the entire population is weak.[157] [158] In those with impaired glucose tolerance or impaired fasting glucose, landmark trials from China,[159] Finland[160] and USA[161] reduced incident cases of type 2 diabetes by up to 33%, 50% and 58% respectively via lifestyle changes (increased exercise, weight loss) and/or pharmacotherapy, though changes may be more modest in a non-trial population. The evidence base for interventions that reliably reduce risk (and therefore incidence) of type 2 diabetes in otherwise healthy adults with normal glycaemic indices is very sparse. At the population level for individuals without diabetes (who may or may not have abnormal glycaemic indices) research indicates that the more behavioural goals that can be attained over time (controlling weight, diet and physical activity) the lower the incidence of type 2 diabetes in the long term.[162] Yet, despite the emerging ability to quantify diabetes risk, there is at present a lack of evidence about how to reduce incidence of diabetes in those at risk of diabetes (as oppose to those with pre-diabetes). More research is needed in this area. For example, a large multicentre randomised control trial comparing different interventions (e.g. lifestyle changes, metformin, and placebo).

172

Using small area maps to plot risk of chronic disease at a local level is relatively novel. It informs visualisation of important social determinants of health which may generate engagement of people with an interest (including local populations) in research and targeted initiatives for improvement. However, the use of this technique beyond the research environment may be limited by governance and technical factors and by the specialist skills needed for the data extraction and mapping. The methodology could be refined through further research of potential utility, to improve geospatial mapping for public health planning. Further studies of feasibility, impact and cost are needed, as are published information governance guidance on how to handle patient level data for geospatial mapping.

**Reflections on the thesis**

When I approached Professor Greenhalgh and expressed a desire to improve my academic skills in her department, I could not have imagined that so much would have been achieved.

I learned a great deal about systematic review methodology, both quantitative and qualitative, and was able to learn how an intractably complex area, which started with a search which generated almost 15,000 research papers, could be distilled into one key table of readily usable risk scores, ready for application in healthcare. Systematic reviewing is not without its difficulties. Much is dependent on the decisions of the researchers, and I was fortunate to be working with a Professor who demanded a high standard of research with appropriate checks and balances. It was frustrating, if something was changed in the data extraction to have to have a second researcher double check the change, but taught me about the importance of striving for the highest standards in academic research.

The cross-sectional study taught me a lot about the management of enormous databases, how complex they are, the degree of error that is inherent when hundreds of thousands of records are being extracted, cleaned and analysed, and the potential such analysis has for informing population wide interventions. Having firsthand experience of working with such a large database has given me an understanding of the daily workload of data analysts and the time it takes to extract and clean data. Previously when I worked at the Department

of Health I took this for granted; now I can appreciate how much background work is required.

Geospatial mapping in an academic context was a new experience for me, and I learned a lot about using Excel, GIS and Adobe Illustrator.  The time involved to manipulate data and map it at first surprised me and has again given me an appreciation of how asking for a map of certain data has vastly more to it that just hitting an imaginary 'map it' button on a computer.  Handling postcode level information was complex, and at times dealing with the governance issues surrounding extraction, handling, and mapping was stressful.  I am grateful to colleagues and the NIGB for their advice and assistance in this process.  I erred on the side of caution with these issues, and it has taught me a great deal about data protection, which I am already applying within my NHS practice.

The media experience was a bonus to the overall research, and taught me much about how the press report research.  It has whetted my appetite for more work of that nature in the future, although I recognise the risk that a misplaced word or phrase can do.

In summary I enjoyed this research very much, was able to work and learn from highly skilled and able colleagues, and am in a position to now apply the findings in the wider NHS.

**Appendix 1: Systematic review study protocol**

| STUDY PROTOCOL for Noble et al: Systematic review on diabetes risk scores (BMJ 2011/889725) | |
|---|---|
| **Objectives** | |
| Objectives | 1. To summarise characteristics and statistical properties of diabetes risk scores.<br>2. To evaluate the impact of such scores in improving patient-relevant outcomes |
| **Methods: Criteria for considering studies for this review:** | |
| Eligibility criteria for study design | For development and validation of risk scores: prospective cohort studies. For impact studies: any design which illuminates (qualitative) or measures (quantitative) the impact of a score on a patient-relevant outcome. |
| Eligibility criteria for patient population | Adults over 18 without diabetes at baseline. |
| Specification of primary endpoint (including measurement instrument) | Development of type 2 diabetes by any measurement technique. |
| Details of subgroups | No subgroups for analysis were predefined in the study protocol. |
| **Methods: Search methods:** | |
| Identification of studies | See Appendix submitted with main paper (briefly, search by experienced librarian and researcher of MEDLINE, Pre-MEDLINE, EMBASE + Cochrane) plus reference search. |
| Efforts to identify ongoing studies | Citation track of all included papers in Google Scholar by two researchers independently. |
| Efforts to identify foreign-language studies | All foreign-language papers deemed relevant according to inclusion and exclusion criteria from the main search were translated by bilingual academics assisted by one of the research team. |
| **Methods: Data collection and analysis:** | |
| List of included studies | See Table 1a, 1b and 1c in supplementary material. |
| Analysis method | Heterogeneity of primary studies precluded formal statistical meta-analysis.<br>Quantitative data were presented in disaggregated form and simple descriptive statistics (e.g. median/range) used to highlight patterns in the data across studies.<br>Qualitative data were analysed using realist methodology. |
| **Management and co-ordination of study:** | |
| Management structure of research team | Core research group chaired by DN and included TG, RM, TD and CM. |

| Data management and quality assurance | Primary extraction of quantitative data was double checked by a second researcher.<br>A one third sample of the qualitative data was double checked by a second researcher.<br>Disagreements were resolved by discussion and at team meetings. |
|---|---|
| Responsibility for statistical analysis | DN and TD.  All presented statistics were discussed at research team meetings. |
| **Publication policy:** | |
| Criteria for authorship | All researchers meeting BMJ criteria for authorship. |
| Writing of paper | See 'contributorship' statement in main paper. |

**Appendix 2: MOOSE checklist**

| Addressed | Criterion | Brief description of how the criteria were handled in the review |
|---|---|---|
| **Reporting of background** | | |
| √ | Problem definition | Diabetes risk scores have existed in the literature for almost 20 years, yet there is confusion amongst GPs and commissioners about [a] the usefulness of the scores [b] which score to use.  We therefore systematically reviewed diabetes risk scores to assess their performance and impact. |
| √ | Hypothesis to be tested | [a] There exist diabetes risk scores which are sufficiently sensitive, specific and discriminatory to have a significant potential impact on patient outcome and [b] such scores have actually had such an impact. |
| √ | Description of study outcomes | Performance as assessed by discrimination, calibration, generalisability and external validation.<br>Impact as assessed by citation tracking and realist review. |
| √ | Type of exposure | Not applicable (risk score validation studies are not designed to assess exposure). |
| √ | Type of study designs used | For development and validation of risk scores: prospective cohort studies.  For impact studies: any design which illuminates (qualitative) or measures (quantitative) the impact of a score on a patient-relevant outcome. |
| √ | Study population | Adults over 18 free of diabetes at baseline. Detailed demography and outcomes extracted and reported in main paper and in detailed Tables 1a, 1b + 1c. |
| **Reporting of search strategy should include** | | |
| √ | Qualifications of searchers | **Helen Elwell** MSc Experienced Librarian at BMA Library, Douglas Noble BM BCh MPH, Catherine Meads previously of NICE Systematic Review Team.  Trish Greenhalgh MD, Professor. |
| √ | Search strategy, including time period included in the synthesis and keywords | Time period: from inception of MEDLINE, EMBASE and Cochrane Library to February 11th 2011.<br>Search strategy – see submitted Appendix. |

| | | |
|---|---|---|
| √ | Databases and registries searched | MEDLINE, Pre-MEDLINE, EMBASE and Cochrane Library. |
| √ | Search software used, name and version, including special features | Ovid was used to search MEDLINE, Pre-Medline and EMBASE.<br>ENDNOTE was used to remove duplicates. |
| √ | Use of hand searching | We searched references of final included papers and other key references. |
| √ | List of citations located and those excluded, including justifications. | Citations were excluded according to exclusion criteria on p8 of the submitted manuscript. Citations describing impact were also reviewed. |
| √ | Method of addressing articles published in languages other than English | We placed no restrictions on language; bilingual academic colleagues translated relevant sections of included papers in dialogue with one of the research team who completed data extraction forms. |
| √ | Method of handling abstracts and unpublished studies | Some studies were excluded on basis of abstract review alone according to exclusion criteria.<br>Unpublished studies were excluded. |
| √ | Description of any contact with authors | Selected authors were contacted as needed to clarify details of study or enquire if further publications. |
| **Reporting of methods should include** | | |
| √ | Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested | Table of included studies 1a, 1b + 1c submitted. |
| √ | Rationale for the selection and coding of data | We extracted data from regression models reporting risk factors for developing type 2 diabetes. This was agreed by the project team. See Tables 1a, 1b + 1c submitted. |
| √ | Assessment of confounding | Confounding is discussed on p17 of the submitted manuscript. |
| √ | Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results | Studies which met performance criteria with regard to discrimination, calibration, generalisability and external validation were independently double-checked by two researchers and agreed by the project team. Meta-analysis was not possible owing to statistical heterogeneity as agreed by the project team and after consultation with other statistical experts. |

| | | |
|---|---|---|
| √ | Assessment of heterogeneity | Gross heterogeneity precluded meta-analysis. |
| √ | Description of statistical methods in sufficient detail to be replicated | Descriptive statistics presented in detail in Tables 1a, 1b and 1c. |
| √ | Provision of appropriate tables and graphics | See Figure 1+2, Table 1 +2, Tables 1a, 1b and 1c. |
| **Reporting of results should include** | | |
| √ | Graph summarizing individual study estimates and overall estimate | Individual risk factors for each diabetes risk score presented in Tables 1a, 1b + 1c.  Also described on p13 of submitted manuscript. |
| √ | Table giving descriptive information for each study included | See Table 1+2, Tables 1a, 1b and 1c. |
| √ | Results of sensitivity testing | Not applicable as meta-analysis precluded. |
| √ | Indication of statistical uncertainty of findings | Addressed in results and discussion section of submitted manuscript. |
| **Reporting of discussion should include** | | |
| √ | Quantitative assessment of bias | Discussed on p17 of submitted manuscript. |
| √ | Justification for exclusion | All studies were excluded based on the pre-defined inclusion criteria. |
| √ | Assessment of quality of included studies | Discussed on p18 of submitted manuscript. |
| **Reporting of conclusions should include** | | |
| √ | Consideration of alternative explanations for observed results | A broad discussion of limitations, impact and future use of risk scores is included in the discussion of submitted manuscript. |
| √ | Generalization of the conclusions | Conclusions are linked to impact and further research is suggested in several areas. |
| √ | Guidelines for future research | See p21 of submitted manuscript. |
| √ | Disclosure of funding source | See statement on p28 of submitted manuscript. |

**Appendix 3: Search strategy for systematic review**

**Database: Ovid MEDLINE(R) <1948 to February week 1 2011> Search Strategy:**

--------------------------------------------------------------------------------
1    diabetes mellitus, type 2/ or prediabetic state/ (64110)
2    diabetes.tw. (233161)
3    ("type 2" or type two or type ii or type II).tw. (131044)
4    2 and 3 (46990)
5    1 or 4 (77482)
6    prediabetic state.tw. (208)
7    pre-diabetic.tw. (396)
8    6 or 7 (602)
9    5 or 8 (77756)
10    odds ratio/ or exp risk/ or regression analysis/ or "sensitivity and specificity"/ or roc curve/ (935102)
11    (risk adj3 (score$ or predict$ or factor$ or model$ or assess$ or calculat$ or analys$ or screen$)).tw. (298546)
12    10 or 11 (1044213)
13    9 and 12 (19631)
14    limit 13 to humans (19334)
15    limit 14 to (classical article or clinical trial, all or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or clinical trial or comparative study or controlled clinical trial or "corrected and republished article" or evaluation studies or introductory journal article or journal article or meta analysis or multicenter study or randomized controlled trial or "review" or technical report or validation studies) (18375)
16    *diabetes mellitus, type 2/ or *prediabetic state/ (49993)
17    15 and 16 (12176)
18    11 and 17 (6169)
***************************

**Database: EMBASE <1980 to 2011 Week 05> Search Strategy:**

--------------------------------------------------------------------------------

1    non insulin dependent diabetes mellitus/ (92300)
2    diabetes.tw. (300352)
3    ("type 2" or type two or type ii or type II).tw. (161037)
4    2 and 3 (66250)
5    1 or 4 (108295)
6    non-insulin dependent.tw. (11795)
7    (prediabetic state or pre-diabetic).tw. (793)
8    6 or 7 (12565)
9    5 or 8 (111058)
10   RISK ASSESSMENT/ or RISK FACTOR/ or RISK/ (732204)
11   LOGISTIC REGRESSION ANALYSIS/ (33553)
12   "sensitivity and specificity"/ (133377)
13   receiver operating characteristic/ (14334)
14   (risk adj3 (score$ or predict$ or factor$ or model$ or assess$ or calculat$ or analys$ or screen$)).tw. (381302)
15   10 or 11 or 12 or 13 or 14 (1013026)
16   9 and 15 (25922)
17   limit 16 to human (22923)
18   14 and 17 (12551)
19   *non insulin dependent diabetes mellitus/ (58830)
20   18 and 19 (6947)
***********************

**Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <February 09, 2011> Search Strategy:**

--------------------------------------------------------------------------------

1    diabetes.tw. (10421)
2    ("type 2" or type two or type ii or type II).tw. (5982)
3    prediabetic state.tw. (7)
4    pre-diabetic.tw. (22)
5    (risk adj3 (score$ or predict$ or factor$ or model$ or assess$ or calculat$ or analys$ or screen$)).tw. (14580)
6    1 and 2 (3254)
7    3 or 4 or 6 (3271)
8    5 and 7 (524)
**************************

**COCHRANE LIBRARY**
**Search History:**

| #1 | MeSH descriptor Diabetes Mellitus, Type 2, this term only | 6547 |
| #2 | MeSH descriptor    Risk explode all trees | 23562 |
| #3 | (#1 AND #2) | 716 |

**Appendix 4: BMJ paper of systematic review**

Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. *BMJ* 2011;343:d7163.

Can be accessed online at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3225074/ and http://www.bmj.com/content/343/bmj.d7163

**Appendix 5: Submitted BJGP paper of cross-sectional study**

This paper is now in press, and will contain some additional material with updates and minor corrections:

Mathur R, Noble D, Smith D, Greenhalgh T, Robson J. Quantifying risk of type 2 diabetes in East London using the QDScore. British Journal of General Practice. *In press.*

**Appendix 6: BMJ Open paper of geospatial mapping study**

Noble D, Smith D, Mathur R, Robson J, Greenhalgh T. Feasibility study of geospatial mapping of chronic disease risk to inform public health commissioning. *BMJ Open* 2012;2(1):e000711.

Can be accessed online at:

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3282296/?tool=pubmed and

http://bmjopen.bmj.com/content/2/1/e000711.full

**Appendix 7: Submitted BMC Public Health paper of mapping and information governance**


This paper was peer reviewed by BMC Public Health.  It is currently being revised, and the plan is for it to be submitted to a different journal.

# References

1. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes research and clinical practice* 2010;87(1):4-14.

2. WHO Diabetes Fact sheet N°312 August 2011. http://www.who.int/mediacentre/factsheets/fs312/en/ Accessed 4th June 2012.

3. Towards implementation of UN General Assembly resolution A/RES/64/265 "Prevention and control of non-communicable diseases" WHO Information Note - Issue 1, Geneva, 23 July 2010.

4. Cerqueira MT, Cravioto A, Dianis N, Ghannem H, Levitt N, Yan LL, et al. Global response to non-communicable disease. *BMJ* 2011;342:d3823.

5. Holman N, Forouhi NG, Goyder E, Wild SH. The Association of Public Health Observatories (APHO) Diabetes Prevalence Model: estimates of total diabetes prevalence for England, 2010-2030. *Diabetic medicine : a journal of the British Diabetic Association* 2011;28(5):575-82.

6. National Diabetes, Clearinghouse. National Diabetes Statistics 2011. Bethesda, Maryland. National Institute of Health, USA. http://diabetes.niddk.nih.gov/dm/pubs/statistics/#Diagnosed20 Accessed 4th July 2012.

7. Yang SH, Dou KF, Song WJ. Prevalence of diabetes among men and women in China. *N Engl J Med* 2010;362(25):2425-26.

8. Saadi H, Carruthers SG, Nagelkerke N, Al-Maskari F, Afandi B, Reed R, et al. Prevalence of diabetes mellitus and its complications in a population-based sample in Al Ain, United Arab Emirates. *Diabetes research and clinical practice* 2007;78(3):369-77.

9. Williamson DF, Vinicor F, Bowman BA. Primary prevention of type 2 diabetes mellitus by lifestyle intervention: implications for health policy. *Ann Intern Med* 2004;140(11):951-57.

10. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004;27(5):1047-53.

11. Diabetes in the UK 2010: Key statistics on diabetes http://www.diabetes.org.uk/Documents/Reports/Diabetes_in_the_UK_2010.pdf Accessed 4th June 2012.

12. Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. *J Epidemiol Community Health* 2009;63(4):332-6.

13. National Institute for Health and Clinical Excellence: Public Health Guidance 35. Preventing type 2 diabetes: population and community interventions.  May 2011. http://guidance.nice.org.uk/PH35 Accessed 20th Feb 2012.

14. National diabetes audit executive summary 2008/9. London: NHS National Information Centre 2009.

15. Banerji MA, Faridi N, Atluri R, Chaiken RL, Lebovitz HE. Body composition, visceral fat, leptin, and insulin resistance in Asian Indian men. *J Clin Endocrinol Metab* 1999;84(1):137-44.
16. McKeigue PM, Shah B, Marmot MG. Relation of central obesity and insulin resistance with high diabetes prevalence and cardiovascular risk in South Asians. *Lancet* 1991;337(8738):382-6.
17. McKeigue PM, Pierpoint T, Ferrie JE, Marmot MG. Relationship of glucose intolerance and hyperinsulinaemia to body fat pattern in south Asians and Europeans. *Diabetologia* 1992;35(8):785-91.
18. Turning the Corner: Improving Diabetes Care.  Department of Health http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/docume nts/digitalasset/dh_4136011.pdf, 2006.  Accessed 4th June 2012.
19. King P, Peacock I, Donnelly R. The UK prospective diabetes study (UKPDS): clinical and therapeutic implications for type 2 diabetes. *Br J Clin Pharmacol* 1999;48(5):643-8.
20. Harris MI, Klein R, Welborn TA, Knuiman MW. Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis. *Diabetes Care* 1992;15(7):815-9.
21. Yorkshire and Humber Public Health Observatory http://www.yhpho.org.uk/default.aspx?RID=81090 Accessed June 4th 2012.
22. Clinical Effectiveness Group Annual Report 2010-2011. http://www.icms.qmul.ac.uk/chs/ceg/annual_report/index.html Accessed June 4th 2012.
23. The London Borough of Newham Joint Strategic Needs Assessment 2010 http://www.newham.gov.uk/NR/rdonlyres/A4B9FD10-CC17-491E-8E93-52AF25DE2E25/0/NewhamJSNA2010.pdf Accessed June 4th 2012.
24. City and Hackney Health and Wellbeing Profile 2010/2011 (Joint Strategic Needs Assessment) http://www.cityandhackney.nhs.uk/pdf/City%20and%20Hackney%20Health%20an d%20Wellbeing%20Profile%202010-11v1.pdf Accessed June 4th 2012.
25. National Survey of People with Diabetes 2006: Picker Institute Europe, 2007.
26. Tower Hamlets Joint Strategic Needs Assessment 2010-2011 http://www.towerhamlets.gov.uk/lgsl/701-750/732_jsna.aspx Accessed 4th June 2012.
27. Simmons RK, Unwin N, Griffin SJ. International Diabetes Federation: An update of the evidence concerning the prevention of type 2 diabetes. *Diabetes research and clinical practice* 2010;87(2):143-49.
28. Yates T, Davies M, Khunti K. Preventing type 2 diabetes: can we make the evidence work? *Postgrad Med J* 2009;85(1007):475-80.
29. National Institute for Health and Clinical Excellence. Lipid modification. 2011. www.nice.org.uk/nicemedia/live/11982/40675/40675.pdf Accessed June 5th 2012.
30. McElduff P, Lyratzopoulos G, Edwards R, Heller RF, Shekelle P, Roland M. Will changes in primary care improve health outcomes? Modelling the impact of financial incentives introduced to improve quality of care in the UK. *Qual Saf Health Care* 2004;13(3):191-97.

31. Griffin SJ, Little PS, Hales CN, Kinmonth AL, Wareham NJ. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. *Diabetes Metab Res Rev* 2000;16(3):164-71.
32. Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med* 2002;136(8):575-81.
33. Lindstrom J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care* 2003;26(3):725-31.
34. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ* 2009;338:b604.
35. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375.
36. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;338:b606.
37. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605.
38. Framingham Heart Study http://www.framinghamheartstudy.org/, Accessed 5th May 2011.
39. Aekplakorn W, Bunnag P, Woodward M, Sritara P, Cheepudomwit S, Yamwong S, et al. A risk score for predicting incident diabetes in the Thai population. *Diabetes Care* 2006;29(8):1872-77.
40. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27(2):157-72; discussion 207-12.
41. Wright C, Dent T. *Quality standards in risk prediction: summary of an expert meeting. http://www.phgfoundation.org* Cambridge: PHG Foundation, 2011.
42. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336(7659):1475-82.
43. NHS Health Check Programme 2009. http://www.healthcheck.nhs.uk/ Accessed 24th Feb 2012.
44. Dent TH, Wright CF, Stephan BC, Brayne C, Janssens AC. Risk prediction models: a framework for assessment. *Public Health Genomics* 2011;15(2):98-105.
45. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med* 2009;150(11):795-802.
46. Pawson R, Tilley N. *Realistic evaluation*. London: Sage, 1997.
47. Kahn KS, Kunz.R, Kleijnen J, Antes G. *Systematic Reviews to Support Evidence-Based Medicine, 2nd Edition*. London: Royal Society of Medicine, 2005.
48. Centre for R, Dissemination. *Systematic Reviews: Centre for Reviews and Dissemination guidance for undertaking reviews in health care*. York: York Publishing Services. Available on http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf, 2009.

49. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. *J.Health Serv.Res.Policy* 2005;10 Suppl 1:21-34.

50. Balkau B, Lange C, Fezeu L, Tichet J, de Lauzon-Guillain B, Czernichow S, et al. Predicting diabetes: clinical, biological, and genetic approaches: data from the Epidemiological Study on the Insulin Resistance Syndrome (DESIR). *Diabetes Care* 2008;31(10):2056-61.

51. Chen L, Magliano DJ, Balkau B, Colagiuri S, Zimmet PZ, Tonkin AM, et al. AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. *Med J Aust* 2010;192(4):197-202.

52. Chuang SY, Yeh WT, Wu YL, Chang HY, Pan WH, Tsao CK. Prediction equations and point system derived from large-scale health check-up data for estimating diabetic risk in the Chinese population of Taiwan. *Diabetes research and clinical practice* 2011;92(1):128-36.

53. Gao WG, Qiao Q, Pitkaniemi J, Wild S, Magliano D, Shaw J, et al. Risk prediction models for the development of diabetes in Mauritian Indians. *Diabet.Med* 2009;26(10):996-1002.

54. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. *BMJ* 2009;338:b880.

55. Joseph J, Svartberg J, Njolstad I, Schirmer H. Incidence of and risk factors for type-2 diabetes in a general population: the Tromso Study. *Scand.J Public Health* 2010;38(7):768-75.

56. Kahn HS, Cheng YJ, Thompson TJ, Imperatore G, Gregg EW. Two risk-scoring systems for predicting incident diabetes mellitus in U.S. adults age 45 to 64 years. *Ann Intern Med* 2009;150(11):741-51.

57. Kolberg JA, Jorgensen T, Gerwien RW, Hamren S, McKenna MP, Moler E, et al. Development of a type 2 diabetes risk model from a panel of serum biomarkers from the Inter99 cohort. 7 ed. United States: Tethys Bioscience, Emeryville, California, USA. jkolberg@tethysbio.com, 2009:1207-12.

58. Liu M, Pan C, Jin M. A Chinese diabetes risk score for screening of undiagnosed diabetes and abnormal glucose tolerance. *Diabetes Technol.Ther.* 2011;13(5):501-07.

59. Mehrabi Y, Sarbakhsh P, Hadaegh F, Khadem-Maboudi A. Prediction of diabetes using logic regression. *Iranian Journal of Endocrinology & Metabolism* 2010;12(1):16-24.

60. Rathmann W, Kowall B, Heier M, Herder C, Holle R, Thorand B, et al. Prediction models for incident type 2 diabetes mellitusin the older population: KORA S4/F4 cohort study. *Diabet.Med* 2010;27(10):1116-23.

61. Schmidt MI, Duncan BB, Bang H, Pankow JS, Ballantyne CM, Golden SH, et al. Identifying individuals at high risk for diabetes: The Atherosclerosis Risk in Communities study. *Diabetes Care* 2005;28(8):2013-18.

62. Schulze MB, Weikert C, Pischon T, Bergmann MM, Al-Hasani H, Schleicher E, et al. Use of multiple metabolic and genetic markers to improve the prediction of type 2 diabetes: the EPIC-Potsdam Study. *Diabetes Care* 2009;32(11):2116-19.

63. Stern MP, Morales PA, Valdez RA, Monterrosa A, Haffner SM, Mitchell BD, et al. Predicting diabetes. Moving beyond impaired glucose tolerance. *Diabetes* 1993;42(5):706-14.

64. Von EA, Schulte H, Assmann G. Risk for diabetes mellitus in middle-aged Caucasian male participants of the PROCAM study: implications for the definition of impaired fasting glucose by the American Diabetes Association. Prospective Cardiovascular Munster. *J Clin Endocrinol.Metab* 2000;85(9):3101-08.

65. Wannamethee SG, Papacosta O, Whincup PH, Thomas MC, Carson C, Lawlor DA, et al. The potential for a two-stage diabetes risk algorithm combining non-laboratory-based scores with subsequent routine non-fasting blood tests: results from prospective studies in older men and women. *Diabet.Med* 2011;28(1):23-30.

66. Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB, Sr. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med* 2007;167(10):1068-74.

67. Alssema M, Feskens EJ, Bakker SJ, Gansevoort RT, Boer JM, Heine RJ, et al. [Finnish questionnaire reasonably good predictor of the incidence of diabetes in The Netherlands]. *Ned.Tijdschr.Geneeskd.* 2008;152(44):2418-24.

68. Alssema M, Vistisen D, Heymans MW, Nijpels G, Glumer C, Zimmet PZ, et al. The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes. *Diabetologia* 2011;54(5):1004-12.

69. Bozorgmanesh M, Hadaegh F, Azizi F. Transportability of the updated diabetes prediction model from Atherosclerosis Risk in Communities Study to a Middle Eastern adult population: community-based cohort study. *Acta Diabetol.* 2010.

70. Bozorgmanesh M, Hadaegh F, Zabetian A, Azizi F. San Antonio heart study diabetes prediction model applicable to a Middle Eastern population? Tehran glucose and lipid study. *Int J Public Health* 2010;55(4):315-23.

71. Cameron AJ, Magliano DJ, Zimmet PZ, Welborn TA, Colagiuri S, Tonkin AM, et al. The metabolic syndrome as a tool for predicting future diabetes: the AusDiab study. *J Intern Med* 2008;264(2):177-86.

72. Collins GS, Altman DG. External validation of QDSCORE((R)) for predicting the 10-year risk of developing Type 2 diabetes. *Diabet.Med* 2011;28(5):599-607.

73. Guerrero-Romero F, Rodriguez-Moran M. [Validation of an instrument for screening cases of type 2 diabetes and monitoring at-risk individuals in Mexico]. *Rev Panam.Salud Publica* 2010;27(3):181-86.

74. Kanaya AM, Wassel Fyr CL, de RN, Shorr RI, Schwartz AV, Goodpaster BH, et al. Predicting the development of diabetes in older adults: the derivation and validation of a prediction rule. *Diabetes Care* 2005;28(2):404-08.

75. Mainous AG, III, Diaz VA, Everett CJ. Assessing risk for development of diabetes in young adults. *Ann Fam Med* 2007;5(5):425-29.

76. Mann DM, Bertoni AG, Shimbo D, Carnethon MR, Chen H, Jenny NS, et al. Comparative validity of 3 diabetes mellitus risk prediction scoring models in a multiethnic US cohort: the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol* 2010;171(9):980-88.
77. McNeely MJ, Boyko EJ, Leonetti DL, Kahn SE, Fujimoto WY. Comparison of a clinical model, the oral glucose tolerance test, and fasting glucose for prediction of type 2 diabetes risk in Japanese Americans. *Diabetes Care* 2003;26(3):758-63.
78. Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 2008;359(21):2208-19.
79. Nichols GA, Brown JB. Validating the Framingham Offspring Study equations for predicting incident diabetes mellitus. *Am J Manag Care* 2008;14(9):574-80.
80. Rahman M, Simmons RK, Harding AH, Wareham NJ, Griffin SJ. A simple risk score identifies individuals at high risk of developing Type 2 diabetes: a prospective cohort study. *Fam Pract* 2008;25(3):191-96.
81. Talmud PJ, Hingorani AD, Cooper JA, Marmot MG, Brunner EJ, Kumari M, et al. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. *BMJ* 2010;340:b4838.
82. Urdea M, Kolberg J, Wilber J, Gerwien R, Moler E, Rowe M, et al. Validation of a multimarker model for assessing risk of type 2 diabetes from a five-year prospective study of 6784 Danish people (Inter99). *J Diabetes Sci Technol.* 2009;3(4):748-55.
83. Wannamethee SG, Shaper AG, Lennon L, Morris RW. Metabolic syndrome vs Framingham Risk Score for prediction of coronary heart disease, stroke, and type 2 diabetes mellitus. *Arch Intern Med* 2005;165(22):2644-50.
84. Bozorgmanesh M, Hadaegh F, Ghaffari S, Harati H, Azizi F. A simple risk score effectively predicted type 2 diabetes in Iranian adult population: population-based cohort study. *Eur J Public Health* 2011.
85. Chien K, Cai T, Hsu H, Su T, Chang W, Chen M, et al. A prediction model for type 2 diabetes risk among Chinese people. *Diabetologia* 2009;52(3):443-50.
86. Rosella LC, Manuel DG, Burchill C, Stukel TA. A population-based risk algorithm for the development of diabetes: development and validation of the Diabetes Population Risk Tool (DPoRT). *J Epidemiol Community Health* 2011.
87. Schulze MB, Hoffmann K, Boeing H, Linseisen J, Rohrmann S, Mohlig M, et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. *Diabetes Care* 2007;30(3):510-15.
88. Simmons RK, Harding AH, Wareham NJ, Griffin SJ. Do simple questions about diet and physical activity help to identify those at risk of Type 2 diabetes? *Diabet.Med* 2007;24(8):830-35.
89. Sun F, Tao Q, Zhan S. An accurate risk score for estimation 5-year risk of type 2 diabetes based on a health screening population in Taiwan. *Diabetes research and clinical practice* 2009;85(2):228-34.
90. Colagiuri S, Vita P, Cardona-Morrell M, Singh MF, Farrell L, Milat A, et al. The Sydney Diabetes Prevention Program: a community-based translational study. *BMC Public Health* 2010;10:328.

91. Prevention of Type 2 Diabetes Program.  Australian type 2 diabetes risk assessment tool (AUSDRISK). http://www.health.gov.au/internet/main/publishing.nsf/Content/chronic-diab-prev-aus Accessed 19th May 2012.

92. Pasco JA, Kotowicz MA, Henry MJ, Nicholson GC. Evaluating AUSDRISK for predicting incident diabetes in an independent sample of women. *Med J Aust* 2010;193(6):374.

93. Harrell F. *Regression Modelling Strategies with Applications to Linear Models, Logistic Regression and Survival Analysis*. New York: Springer-Verlag, 2001.

94. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med* 2011;9:103.

95. Lindstrom J, Absetz P, Hemio K, Peltomaki P, Peltonen M. Reducing the risk of type 2 diabetes with nutrition and physical activity - efficacy and implementation of lifestyle interventions in Finland. *Public Health Nutr.* 2010;13(6A):993-99.

96. Buijsse B, Simmons RK, Griffin SJ, Schulze MB. Risk assessment tools for identifying individuals at risk of developing type 2 diabetes. *Epidemiol Rev* 2011;33(1):46-62.

97. Collins G, Altman D. Predicting the risk of chronic kidney disease in the UK: an evaluation of QKidney(R) scores using a primary care database. *Br J Gen Pract* 2012;62(597):243-50.

98. Collins GS, Altman DG. An independent and external validation of QRISK2 cardiovascular disease risk score: a prospective open cohort study. *BMJ* 2010;340:c2442.

99. Collins GS, Altman DG. An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. *BMJ* 2009;339:b2584.

100. Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. *BMJ* 2011;342:d3651.

101. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989;13(3):319-40.

102. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. Diffusion of innovations in service organisations: systematic literature review and recommendations for future research. *Millbank Quarterly* 2004;82:581-629.

103. Wareham NJ, Griffin SJ. Risk scores for predicting type 2 diabetes: comparing axes and spades. *Diabetologia* 2011;54(5):994-95.

104. Herman WH. Predicting risk for diabetes: choosing (or building) the right model. *Ann Intern Med* 2009;150(11):812-14.

105. Department for Communities and Local Government Indices of Deprivation 2007 http://www.communities.gov.uk/publications/communities/indiciesdeprivation07 Accessed June 6th 2012.

106. The English Indices of Deprivation 2010 http://www.communities.gov.uk/publications/corporate/statistics/indices2010 Accessed June 6th 2012.

107. Mathematical equation for calculating QDScore
http://svn.clinrisk.co.uk/opensource/qdscore/ Accessed 6th June 2012.
108. *Stata Statistical Software:  Release 10* [program]. College Station, TX: StataCorp LP, 2007.
109. Clinical Effectiveness Group.  Diabetes Risk 2011.  Queen Mary University London.  ISBN: 978-0-9567899-5-2.  http://www.icms.qmul.ac.uk/chs/Docs/63543.pdf Accessed 24th Feb 2012.
110. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. *BMJ* 2011;343:d7163.
111. Abduelkarem AR SS, Hammrouni AM, Aldouibi SS, Albraiki WM, El-Shareif HJ Risk calculation of developing type 2 diabetes in Libyan adults. *Pract Diab Int* 2009 26(4):148-51i.
112. Hjellset VT, Bjorge B, Eriksen HR, Hostmark AT. Risk factors for type 2 diabetes among female Pakistani immigrants: the InvaDiab-DEPLAN study on Pakistani immigrant women living in Oslo, Norway. *J Immigr Minor Health* 2011;13(1):101-10.
113. Rao N, Eastwood SV, Jain A, Shah M, Leurent B, Harvey D, et al. Cardiovascular risk assessment of South Asians in a religious setting: a feasibility study. *Int J Clin Pract* 2011;66(3):262-9.
114. O'Flynn N, Potter J. Commissioning health care in England: an evolving story. *Ann Intern Med* 2011;155(7):465-9.
115. The NHS Information Centre http://www.ic.nhs.uk/ accessed 10th October 2011.
116. Competencies for Clinically-led Commissioning, RCGP Centre for Commissioning, September 2011.
117. Commissioning for the future Learning from a simulation of the health system in 2013/14, The King's Fund 2011.
118. Noble D, Smith D, Mathur R, Robson J, Greenhalgh T. Feasibility study of geospatial mapping of chronic disease risk to inform public health commissioning. *BMJ Open* 2012;2(1):e000711.
119. Snow J. *On the mode of communication of cholera*. 2d. ed. London: J. Churchill, 1855.
120. Dorling D. A review of chapter 2, book review article of Tufte E.R. Visual Explanations. *Progress in Human Geography, 23, 1, 127-131* 1999;23, 1:127-31.
121. Stewart JE, Battersby SE, Lopez-De Fede A, Remington KC, Hardin JW, Mayfield-Smith K. Diabetes and the socioeconomic and built environment: geovisualization of disease prevalence and potential contextual associations using ring maps. *Int J Health Geogr* 2011;10:18.
122. Brewer CA. Basic mapping principles for visualizing cancer data using Geographic Information Systems (GIS). *Am J Prev Med* 2006;30(2 Suppl):S25-36.
123. Gilbert E. Pioneer Maps of Health and Disease in England. *The Geographical Journal* 1958;124(2):172-83.
124. Kwan M, Lee J. Spatially Integrated Social Science: Examples in Best Practice. In: Michael F, Donald G, editors: Oxford University Press, 2003.
125. Curtis AJ, Lee WA. Spatial patterns of diabetes related health problems for vulnerable populations in Los Angeles. *Int J Health Geogr* 2010;9:43.

126. CDC Disease Maps 2011 http://www.cdc.gov/obesity/data/trends.html Accessed June 6th 2012.
127. Barker LE, Kirtland KA, Gregg EW, Geiss LS, Thompson TJ. Geographic distribution of diagnosed diabetes in the U.S.: a diabetes belt. *Am J Prev Med* 2011;40(4):434-9.
128. Green C, Hoppa RD, Young TK, Blanchard JF. Geographic analysis of diabetes prevalence in an urban area. *Soc Sci Med* 2003;57(3):551-60.
129. Congdon P. Estimating prevalence of coronary heart disease for small areas using collateral indicators of morbidity. *Int J Environ Res Public Health* 2010;7(1):164-77.
130. Silman A, Harrison B, Barrett E, Symmons D. The existence of geographical clusters of cases of inflammatory polyarthritis in a primary care based register. *Ann Rheum Dis* 2000;59(2):152-4.
131. de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Fam Pract* 2006;23(2):253-63.
132. Quality Outcomes Framework 2010/11 http://www.qof.ic.nhs.uk/ Accessed June 6th 2012.
133. Martin D, Wright JA. Disease prevalence in the English population: a comparison of primary care registers and prevalence models. *Soc Sci Med* 2009;68(2):266-74.
134. Soljak M, Samarasundera E, Indulkar T, Walford H, Majeed A. Variations in cardiovascular disease under-diagnosis in England: national cross-sectional spatial analysis. *BMC Cardiovasc Disord* 2011;11:12.
135. Bulletin 2011/02: ONS Postcode Directory and National Statistics Postcode Look-up, February 2011 http://www.ons.gov.uk/ons/guide-method/geography/news/bulletin-2011-02--ons-postcode-directory-and-national-statistics-postcode-look-up--february-2011/index.html Accessed June 6th 2012.
136. Generalised Land Use Database Statistics for England 2005. Department for Communities and Local Government 2007.
137. Environmental Systems Research Institute, Redlands, California. http://www.esri.com/software/arcgis/index.html Accessed June 6th 2012.
138. Berke O. Exploratory disease mapping: kriging the spatial risk function from regional count data. *Int J Health Geogr* 2004;3(1):18.
139. Charles Booth and the survey into life and labour in London (1886-1903), http://booth.lse.ac.uk/.
140. Greenhalgh T, Stramer K, Bratan T, Byrne E, Russell J, Potts HW. Adoption and non-adoption of a shared electronic summary record in England: a mixed-method case study. *BMJ* 2010;340:c3111.
141. Feng J, Glass TA, Curriero FC, Stewart WF, Schwartz BS. The built environment and obesity: a systematic review of the epidemiologic evidence. *Health Place* 2010;16(2):175-90.
142. Wilkinson R, Pickett K. *The Spirit Level: Why equality is better for everyone*. London: Penguin Books, 2009.
143. Swinburn B, Egger G, Raza F. Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity. *Preventive Medicine* 1999;29(6):563-70.
195

144. Office for National Statistics.  Review of the Dissemination of Health Statistics: Confidentiality Guidance, 2006.  .
145. Association of Public Health Observatories: Using small area data in public health intelligence, August 2009.  http://www.apho.org.uk/resource/item.aspx?RID=74894 Accessed 8th December 2011.
146. Modern health mapping shows how poverty and ill health persist over 100 years, http://www.qmul.ac.uk/media/news/items/smd/63821.html Thursday 16 February 2012.
147. The changing face of poverty: Updated maps of London's poorest areas show epidemic of 'junk food' diabetes in same streets where Victorians died of malnutrition, http://www.dailymail.co.uk/news/article-2102485/Updated-maps-Londons-poorest-areas-epidemic-junk-food-diabetes-streets-Victorians-died-malnutrition.html 17 February 2012.
148. Poverty and ill health still go hand in hand, East End Life 27th Feb 2012.
149. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14(1):32-8.
150. Rose G. Strategy of prevention: lessons from cardiovascular disease. *Br Med J (Clin Res Ed)* 1981;282(6279):1847-51.
151. Hofman A, Vandenbroucke JP. Geoffrey Rose's big idea. *BMJ* 1992;305(6868):1519-20.
152. Fair Society Healthy Lives.  The Marmot Review.  2010. http://www.instituteofhealthequity.org/projects/fair-society-healthy-lives-the-marmot-review Accessed 24th Feb 2012.
153. Political declaration of the High-level Meeting of the General Assembly on the Prevention and Control of Non-communicable Diseases, United Nations General Assembly 16 September 2011.
154. Tackling Obesities: Future Choices – Summary of Key Messages, http://www.bis.gov.uk/assets/bispartners/foresight/docs/obesity/20.pdf Accessed 5th January 2012.
155. Preventing type 2 diabetes: risk identification and interventions for individuals at high risk.  National Institute for Health and Clinical Excellence: public health draft guidance. Nov 2011. http://www.nice.org.uk/guidance/index.jsp?action=download&o=57026 Accessed 24th Feb 2012.
156. Saaristo T, Peltonen M, Keinanen-Kiukaanniemi S, Vanhala M, Saltevo J, Niskanen L, et al. National type 2 diabetes prevention programme in Finland: FIN-D2D. *Int J Circumpolar.Health* 2007;66(2):101-12.
157. Gillies CL, Lambert PC, Abrams KR, Sutton AJ, Cooper NJ, Hsu RT, et al. Different strategies for screening and prevention of type 2 diabetes in adults: cost effectiveness analysis. *BMJ* 2008;336(7654):1180-85.
158. Simmons RK, Echouffo-Tcheugui JB, Griffin SJ. Screening for type 2 diabetes: an update of the evidence. *Diabetes Obes.Metab* 2010;12(10):838-44.
159. Pan XR, Li GW, Hu YH, Wang JX, Yang WY, An ZX, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. *Diabetes Care* 1997;20(4):537-44.

160. Tuomilehto J, Lindstrom J. The major diabetes prevention trials. *Curr.Diab.Rep.* 2003;3(2):115-22.
161. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346(6):393-403.
162. Simmons RK, Harding AH, Jakes RW, Welch A, Wareham NJ, Griffin SJ. How much might achievement of diabetes prevention behaviour goals reduce the incidence of diabetes if implemented at the population level? *Diabetologia* 2006;49(5):905-11.