Coding, Classification and Diagnosis of Diabetes

A review of the coding, classification and diagnosis of diabetes in primary care in England with recommendations for improvement
This report has received Department of Health Gateway clearance number 14664.
# Contents

1. Preface 4

2. Executive Summary 5

3. Introduction 7

4. The patients’ perspective 9

5. Evidence 10
   a. A systematic review 10
   b. Analysis of diagnostic databases 16
   c. Evidence from the pilot studies 24

6. Improving diagnosis 31
   a. Guidelines for classification 31

7. Improving existing patient records 34
   a. Good coding practice 34
   b. Support tools 41

8. Appendices 49
   a. Diabetes diagnosis guidelines 49
   b. Clinical criteria, insulin and diagnosis of diabetes 50
   c. Acknowledgements 51
1. Preface

It used to be thought that diabetes was a simple diagnosis - the patient was either Type 1 or Type 2. Type 1 diabetes usually occurs in younger people with sudden onset, when the pancreas no longer produces insulin and patients have to start insulin therapy straight away. In Type 2 diabetes, the pancreas either fails to make enough insulin or the body fails to respond fully to the insulin that it does produce. People with Type 2 are normally older, have often had the condition for some years and can be treated with diet and exercise, and/or medications and/or insulin therapy. However, in reality it is very often much less clear-cut in making the diagnosis of diabetes. Recent research shows that there are now many variants of diabetes and no doubt many more are yet to be discovered.

People with a health problem want to know what is wrong with them, what it means and what can be done to help them either recover or stay as well as possible whilst living with their condition. Whilst general care can always be provided, a proper diagnosis is needed before any disease or condition can be fully treated or managed effectively.

Diabetes is a complex condition that affects all parts of the body and may not always be easy to diagnose at first presentation. This challenge was previously considered by the former National Clinical Director for Diabetes, Dr Sue Roberts CBE who established a group to explore the evidence and extent of misdiagnosis and to make recommendations to assist General Practitioners in diagnosing and classifying the condition. This report shows that although the newly discovered variants of diabetes do make diagnosis more complex, Type 1 and Type 2 still represent the majority of cases. However, in a significant number of patients, mistakes are being made in identification of the type of diabetes. In particular, some Type 2 patients on insulin therapy are mistakenly labelled as having Type 1 diabetes. Mistakes like this or other errors can impact on patient information, education, treatment and their health outcomes.

This report brings together evidence of the impact an incorrect diagnosis can have on a person with diabetes, describes the current research base for evidence of misdiagnosis and misclassification, and shows the extent of that misclassification and misdiagnosis through an analysis of GPs’ records.

The report provides front line staff with a simple and easy to use classification algorithm to help make more accurate diagnoses.

The days of a simple diagnosis of either Type 1 or Type 2 are over. Of course there will be situations where the type of diabetes is unclear at first diagnosis, but treatment is still available and the diagnosis may become clearer over time. Further tests and specialist advice can help. This should be explained to the patient and the notes coded appropriately.

The contents of this report are a significant step forward in ensuring that when people are diagnosed with diabetes, the diagnosis is right and appropriate treatment is recommended. It will be of interest to anyone involved in providing diabetes care, people with diabetes and researchers.

Our thanks go to all the people who have contributed to this important piece of work.

Dr Rowan Hillson
National Clinical Director for Diabetes

Dr Clare Gerada
RCGP Chair
2. Executive Summary

This chapter summarises the key findings from the report.

These findings have been endorsed by Diabetes UK, Juvenile Diabetes Research Foundation, British Society of Paediatric Endocrinologists and Diabetologists, Primary Care Diabetes Society, Association of British Consultant Diabetologists and the British Computer Society: Primary Healthcare Specialist Group.

The Patient’s perspective
- Misdiagnosis causes increased concern and worry for the people affected
- There is anecdotal evidence that misdiagnosis is widespread
- Misdiagnosis undermines trust in healthcare professionals abilities and judgements.

The Evidence

A systematic review
- Diagnosing diabetes is a complex task especially in the young adult
- There is substantial evidence of the miscoding and misclassification of diabetes
- In the absence of universal access to definitive clinical tests there is a need for pragmatic and relevant clinical classification.

Analysis of diagnostic databases
- 85-90% of data on diabetes is fit for purpose but that there is room for improvement
- Data quality on diagnosis can be improved by comparing it to other data in practice electronic records
- A simple search tool, that can be amended to account for changes in diagnostic criteria, could be embedded in GP computer systems to allow diagnostic review to take place as an integral part of diabetes management in primary care.
- Standardisation of the data collection forms and picking lists would also aid standardisation of data recording
- There are likely to be unmet educational needs about the diagnosis, classification and treatment of diabetes in primary care.

Results of pilot audit
- Investigation of routinely collected clinical data showed 2.2% of people had been mis-diagnosed, 2.1% of people misclassified and 0.9% miscoded
- There may be a link between cases that are not on the QOF register and reduced quality of care
- Six MIQUEST (Morbidity Information Query and Export Syntax) queries were developed to test the face validity of our approach
- An audit tool has been developed to highlight cases that are incorrectly managed.
Improving diagnosis

- Errors in diagnosis of diabetes can be caused by lack of information or understanding by healthcare professionals
- Errors can have a considerable impact on patient care
- It is important to recognise uncertainty in diagnosis in complex cases
- Complexity of diagnosis means there is a lack of clinically based guidelines for the classification of diabetes
- Accurate diagnosis is critical for the appropriate treatment for the person with diabetes
- A proposed guideline algorithm is designed to be pragmatic and easy to use
- The algorithm does not replace expert opinion or aetiological testing
- There will always be cases that do not conform to the proposed classification.

Improving existing patient records

Good coding practice

- The role of general practice is changing from a reactive provider of primary care to a co-ordinator and integrator of care. Computerised medical records can facilitate quality improvement and research, in a way that was not possible with paper records. A high quality record will have sufficient coded data to make it fit for purpose
- Good quality records require ongoing maintenance as diagnostic criteria change and records need to be coded using contemporary classifications and coding systems.

Support tools

- Good quality medical records make it more likely that patients will be managed correctly. In diabetes care, best practice varies according to the type of diabetes
- Correct coding of a number of items underpins our ability to correctly identify cases and provide improved recording of therapy, monitoring and management
- Automated tools have been created that improve data quality in diabetes. These can be found at: http://www.clininf.eu/cod The website contains:
  - User guides
  - MIQUEST queries
  - Excel spreadsheet with embedded macros
  - Useful background reading.

Treatment

Every person with diabetes can be treated. Appropriate classification of the type of diabetes allows optimal treatment to start promptly.
3. Introduction

It is widely recognised that diabetes is one of the most significant threats to the health of people in England. The recent Association of Public Health Observatories (APHO) prevalence model estimates that there are 3.1 million people with diabetes in England\(^1\). This is a 25% increase on previous estimates; with 800,000 of these not diagnosed. Furthermore an ageing population and increasing obesity mean that the number of adults with diabetes is projected to increase significantly over the next twenty years.

By 2020 an estimated 3.8 million adults, or 8.5% of the adult population, will have diabetes and by 2030 this is estimated to rise to 4.6 million or 9.5% of the adult population.\(^1\) Approximately half of this increase is due to the changing age and ethnic group structure of the population and half due to higher levels of obesity.

It is obviously very important that people are accurately diagnosed with diabetes to ensure they receive appropriate healthcare. The genesis of this report lies in the suspicion amongst healthcare professionals that some people identified with diabetes on practice registers might not have the condition. Alternatively, they might have diabetes but not the type their patient record indicated they had. This view was reinforced by the effect of the 2006 Quality and Outcomes Framework (QOF) changes which required that diabetes be reported as Type 1 or Type 2 and saw a 22% reduction in the number of people on diabetes registers.\(^2\) There was also evidence from the 2004 - 5 National Diabetes Audit report which identified that 43% of records did not specify the type of diabetes.\(^3\)

An initial meeting of stakeholders, listed in Appendix C, agreed that there was work to be done on this issue and established a Task and Finish Working Group to take it forward. The Group was broadly based and included experts in the fields of systematic reviews, diagnosing diabetes, clinical data analysis and service users.

However the Group was intent on not just seeing whether or not there were errors but also to establish how widespread the problem was, develop ways of improving future diagnosis and provide support in correcting existing cases where errors had been made.

The first step was to define what errors would be considered and these fell into three categories:

- **Miscoding** is when the wrong computer code is used meaning that it is not possible to determine the type of diabetes precisely.
- **Misclassification** is when someone is incorrectly classified as having a type of diabetes that they don’t have.
- **Misdiagnosis** is when someone is diagnosed with any form of diabetes when they don’t have it at all.

It was on the basis of these definitions that the work was undertaken and it is reported in this document under the following three agreed work programmes themes:

1. **Evidence**: This work programme initially gathered data, both nationally and internationally, on the type and extent of errors in diagnosis of diabetes. This was supplemented by an analysis of nearly one million patient records on diagnostic databases looking for potential errors using a range of clinical, therapeutic and other data. It also includes evidence from a pilot of an audit tool developed to identify potential errors.

2. **Improving diagnosis**: The aim of this work programme was look at how new diagnoses of diabetes can be improved in the future and provides a pragmatic guidance algorithm to support decision making.

3. **Improving existing patient records**: This work programme provides advice and support on how potential errors in primary care records can be identified through use of the audit tool. It also includes advice on good coding practice.
During the progress of the report work led by Professor Andrew Hattersley examined the use of non-invasive techniques to support the classification of diabetes. Their findings on the accuracy of QOF, and other criteria, on classifying insulin dependent Type 1 diabetes are at Appendix B.

References
4. The Patients’ Perspective

This chapter has been written by Avril Surridge and Bob Moberley, NHS Diabetes User Reference Group.

When people are told they have diabetes their initial reaction can be one of shock, alarm, worry or, unfortunately, “so what” as they don’t realise the implications of what they are being told. What they will all believe is that the doctor or nurse has got it right and is giving them accurate information about their diabetes. A major worry is the impact that an incorrect or conflicting diagnosis can have; this can cause even more worry and alarm and undermines people’s confidence in their doctor or nurse’s knowledge of diabetes. It is acknowledged that there is a wide variation in the experience, knowledge and expertise of doctors - as with all other professions - but many people believe in them implicitly!

It concerns people with diabetes greatly that there is apparently no systematic or clinical data relating to patients who have been, or believe they have been, misdiagnosed with diabetes. When people with diabetes are asked whether it is Type 1 or Type 2 common replies we have heard are

“I don’t know. Nobody’s ever told me”

“I can’t remember.”

“What’s the difference?”

Most worryingly people say

“I used to be Type 2 but now I am on insulin so I am Type 1.”

One middle-aged lady said “I’ve had diabetes since I was 12 and I always thought I was a Type 1 but my doctor tells me that now I have reached 50 I’m a Type 2.” Similar statements were repeated across a wide area.

All of this causes many people with diabetes to get very confused about Type 1 and Type 2. Many don’t know or have never had the differences explained to them. Some appear never to have been told if they have Type 1 or 2, and have relied on the word of friends to help them. There are far too many who have Type 2 but when put on to insulin believe that they have then become Type 1. Some of these have been told by a Health Care Professional (HCP) that going on to insulin means that they have become Type 1. Sometimes they are then told, correctly, by another HCP, that this is not the case. This leads to even more confusion and general mistrust of the diabetes knowledge and expertise of the medical profession.

A number of elderly people who were diagnosed with Type 2 about 6 or 7 years ago have never had any medication for it, pay little attention to their diet and never have an HbA1c units 5.2%. They all believe they were diagnosed incorrectly, do not have diabetes and they are probably right.

Getting the right diagnosis of diabetes is the first step towards getting the right treatment not just in terms of prescribing treatments but also in general advice. Increasingly, information for people with diabetes is coming in separate forms for Type 1 and Type 2 and it is important that people are steered to the right information and the right education programme.

If the results of this report are right then there are many people out there who either have been told they have a type of diabetes they do not have or they do not have it at all. It is vital for the potential health outcomes of patients that they are given accurate and trustworthy information from diagnosis. In order to deliver reliability, HCPs need parameters and guidance from those experts who have the necessary knowledge and experience. It is hoped therefore that the results of this project will be accepted and used.
5. Evidence

(a) A systematic review

This section was edited by Professor Kamlesh Khunti, Professor of Primary Care Diabetes and Vascular Medicine, University of Leicester.

Errors and omissions in classifying diabetes

The main content of this chapter has been published as a paper: Incorrect and incomplete coding and classification of diabetes: a systematic review1.

Introduction

The main aim of the project was to carry out a review to identify the types of incorrect or incomplete coding or classification within diabetes or between diabetes and other conditions and the associated implications. The Working Group also considered that it would be useful to look at the evidence about how frequently these errors and omissions happened. The opportunities for misclassification or miscoding are extensive, with the American Diabetes Association describing a complex range of types and sub-types of diabetes including eight broad categories under the heading ‘other types’2, each with further subdivisions including an ‘other’ category. However Type 1 diabetes and Type 2 diabetes are the most common types, with the latter estimated as accounting globally for approximately 90% of diabetes cases3. Additionally there are eight broad categories under the heading ‘other types’, each with further subdivisions. These include latent autoimmune diabetes of adulthood (LADA) and monogenic subcategories including maturity-onset diabetes of the young (MODY). Although the percentage of overall cases of diabetes is low for some classifications, actual numbers are clinically significant and distinguishing by type is vitally important for the individual patients concerned.

The challenges associated with correct classification and administrative coding are increasing as new types are identified and described. Similarities in presentation can also lead to confusion between types; for example, MODY is frequently misdiagnosed as Type 1 because of the young age of onset in non obese people4. In addition, difficulties with distinguishing between types are likely to be exacerbated as patterns of presentation by type change; for example, the increasing incidence of Type 2 in younger people can result in difficulties in differentiating between Type 1 and Type 2 at diagnosis4.

In addition to the number of types of diabetes now recognised, the range of methods of distinguishing between types has also expanded. For example, traditionally, Type 1 and Type 2 have been diagnosed clinically by a fasting plasma glucose test, an oral glucose tolerance test or a random plasma glucose test, combined with an assessment of symptoms. However, newer technologies are also available; for example, determining the presence of islet cell antibodies at diagnosis of diabetes has been shown to improve classification by type5.

There is potential for serious implications resulting from miscoding and misclassification in the context of any diagnosis, particularly where this relates to a chronic condition. These implications may include inappropriate clinical management, negative psychological effects and sub-optimal use of resources. For example, in diabetes, the misdiagnosis of Type 1 as Type 2 may lead to poor management of hyperglycaemia, with other medications being used inappropriately before insulin replacement6. Educational programmes designed for people with specific types of diabetes have been developed and failure to accurately distinguish between
types may mean that patients are not referred to appropriate education programmes7-9. Incorrect coding may have implications for care as patients who have incorrect labels may be missed by searches or fail to benefit from appropriate decision support triggered by a specific code10. In addition to clinical management, it has been shown that labelling can have an extremely important role in the therapeutic process, because patients can relate to a label regardless of their understanding of the illness and their treatment decisions may depend on whether they accept the assigned label11.

Method
No previous published reviews that investigated the broad topic of miscoding and misclassification within diabetes or between diabetes and other conditions were identified. A systematic review was therefore conducted to identify relevant published literature, in order to address our aims as described above.

Results
Combined results from the electronic database searches identified a total of 366 titles and abstracts after removal of duplicates. After matching these against our inclusion and exclusion criteria and using supplementary strategies for obtaining information about potential papers, we identified a total of seventeen papers for data extraction12-28. Six of these reports were from the UK, three from the USA, two each from Germany and from Denmark and one each from France, Spain, the Netherlands and Tanzania.

Types of miscoding or misclassification
The selected papers described a range of instances of incomplete or incorrect coding or classification of diabetes either as the main focus of the paper or an incidental finding. Two of these papers described more than one broad type of miscoding or misclassification23,26. The authors of five papers 13, 21-23, 26 considered the question of accurately distinguishing between Type 1 and Type 2 and in four of these difficulties were described in relation to young people21-23, 26. One of these papers suggested that a new category of diabetes was needed, latent autoimmune diabetes of youth (LADY) for some patients in whom differentiation between Type 1 and Type 2 is problematic21. Three papers provided information about the distinction between diabetes and no-diabetes, including people with limited levels of glycaemia14, 17, 28, and four papers considered the extent and / or implications of incorrect classification relating to MODY18, 20, 24, 27. Two papers described failure to categorise diabetes by type15, 26. The remaining papers described instances of failure to recognise LADA12 pancreatic diabetes18, persistence of foetal haemoglobin19 or initial consideration of a diagnosis of AIDS in patients later identified as having diabetes25.

Quantitative data presented, including extent of miscoding and misclassification
The information provided in the selected papers was mostly descriptive, including ways in which miscoding or misclassification were identified and the implications of these errors, but twelve papers13-18, 20-23, 26, 28 contained some data relating to the extent of these problems. However, the heterogeneity of these studies, even those considering the same broad type of miscoding or misclassification, as well as small samples in many instances, made meta-analysis inappropriate. These considerations also limited the usefulness of drawing any conclusions based on less formal methods of pooling results or of seeking generalisations about the prevalence of incorrect classification or coding. For example, a study which provided information about incorrect coding in the secondary care records of young people in the USA confirmed correct coding in only 16% of cases coded as having Type 2, with the other cases being reclassified as having Type 1 or no diabetes22. This contrasts with a report of fourteen cases of Type 2 in young people which were all correctly classified by paediatricians in the Netherlands in response to a questionnaire asking them to identify new cases of diabetes23.
Lack of clarity in the reporting of some quantitative results was noted; for example, the presentation of findings did not always include details of the exact denominators used for calculating percentages. In a paper considering the distinction between Type 1 or Type 2 and pancreatic diabetes it was unclear whether any of the cases finally assessed as having pancreatic diabetes had been classified as such at baseline. Nevertheless, in spite of the limitations described above, some studies reported high proportions incorrectly classified or coded, confirming that incorrect diagnosis and recording are important problems.

**Implications of miscoding and misclassification**

A number of implications resulting from the types of miscoding or misclassification considered in the review were highlighted by the authors of the selected papers, based either on direct evidence from the studies or speculation about their findings. Implications for clinical management, including treatment options and risk management in patients and their families, were highlighted, as well as financial and psychological consequences and implications related to the validity of quality of care evaluations and research. The sample included only one qualitative study and one case study based on interviews with the patient and his mother, which provided detailed evidence about the psychological implications of receiving an incorrect diagnosis of diabetes type.

**Discussion**

**Overview**

The review identified a number of papers covering a range of different types of incorrect or incomplete coding or classification, particularly in relation to young people. The usefulness of combining data from these studies to estimate the extent of the problem was limited by heterogeneity in terms of the type of misclassification, settings, samples studied and methods used. Nevertheless, the errors or omissions under consideration were identified as occurring with sufficient frequency for the problem to be considered important. The papers included in the final selection highlighted a number of implications for patients, health care providers and others which will be discussed in more detail below. At detailed level, some of these implications are specifically relevant to the distinction between different types of diabetes but, viewed as broad categories, they can also be generalised to overall problems resulting from errors in diagnosis and coding.

**Implications for patients and their families**

Implications for patients highlighted by our review include potential inappropriate or delayed management of their condition. This may be related to pharmacological management options, with patients being put at risk of negative outcomes associated both with medicines that are prescribed inappropriately and with those not prescribed which would improve clinical outcomes. For example, misclassification of LADA as Type 2 diabetes will frequently lead to failure to prescribe insulin at the appropriate stage in the course of hyperglycaemia. Conversely, classification of MODY as Type 1 is likely to result in inappropriate insulin prescribing alongside failure to provide the benefits associated with sulphonylureas in patients with this type of diabetes. For some types of diabetes, specific relevant treatment options may fail to be considered if misclassification occurs, for example pancreatic replacement therapy for people with pancreatic diabetes. In addition, incorrect classification may have implications for the management of risks associated with different types of diabetes; for example, increased genetic risk of having MODY in the families of people with this type of diabetes.
In addition to clinical effects, misclassification associated with diabetes can have serious psychological implications such as those associated with being labelled incorrectly or with lifestyle and overall quality of life. The two studies which were based on interviews also highlighted the fact that negative psychological implications may also affect family members and may persist after the error is identified, for example, in relation to anxiety about stopping insulin and feelings of annoyance about previous inappropriate management. This observation about families would also apply to other carers. A further implication for patients identified by the authors of one of the selected papers is potential financial disadvantage, for example, where people misdiagnosed as having diabetes are prescribed medication which is not required. Though not specifically mentioned in the selected papers, there may also be financial implications associated with occupational disadvantage resulting from an incorrect diabetes-related diagnosis.

**Implications for health care providers, managers and researchers**

The increasing complexity of diabetes classification and of the methods available for distinguishing between types has added to the challenges facing health care providers responsible for making accurate and complete diagnoses within diabetes. The difficulty of differentiating by type has been acknowledged even by experts in the field, for example, in relation to the predictive value of islet cell antibody testing. These challenges have implications for clinical practice relating to treatment options and risk management, as described above.

In addition, incorrect and incomplete coding and classification may have financial implications for health care practitioners, managers and policy makers. For example, one of the selected papers highlighted the possible impact on incentive payments based on the Quality and Outcomes Framework in UK primary care, which may be affected by incorrect identification of cases according to diabetes type. The validity of the results of quality of care measurement may also be affected, with a consequent impact on the usefulness of auditing outcomes and processes in order to improve patient care. Where research involves cases where there has been miscoding or misclassification, there may be implications for the validity of findings, as suggested in relation to incorrect or inconclusive coding by diabetes type in young people.

**Differences between miscoding and misdiagnosis**

In considering the question of disease classification, in terms of the distinction either between types of diabetes or between diabetes and other conditions, papers were included describing both miscoding and misdiagnosis. However, it should be noted that an incorrect or incomplete code does not necessarily have the same implications as an error in diagnosis. An incorrect or incomplete code recorded in a database may or may not reflect failure to make a correct and complete diagnosis. For example, a patient recorded simply as having diabetes, or as having the wrong type of diabetes, may well be recognised by the responsible clinician as having a specific, correct type and the error may be simply administrative. This would not, therefore have implications for current patient management but there is, nevertheless, potential for impact in other areas, for example, in relation to incentive payments, the validity of research or audit and the appropriate allocation of resources which are specific to particular diabetes subgroups.
References


22. Rhodes ET, Laffel LMB, Gonzalez TV, Ludwig DS. Accuracy of Administrative Coding for Type 2 Diabetes in Children, Adolescents, and Young Adults. Diabetes Care 2007; 30: 141-143.


(b) Analysis of Diagnostic Databases

Is there a problem with classifying diabetes?

The contents of this chapter are an edited version of *A method of correcting miscoding, misclassification and misdiagnosis in diabetes: a pilot and validation study of routinely collected data*.

Introduction

Accurate classification of disease is vital for effective disease management, audit of the quality of care and research. The World Health Organisation (WHO) classifies diabetes into Type 1, Type 2 and four other types covering genetic forms, drug or chemical induced, gestational, and unknown. The latter including those that do not fit clearly into any of the other categories. All of these categories have guidelines available to support and inform the management of diabetes; for example the American National Guidelines Clearing House contains 162 diabetes guidelines; in England National Institute of Clinical Evidence and Health (NICE) make separate recommendations for the effective management of diabetes. However all this advice relies on an accurate diagnosis in the first place.

Much of primary care is computerised and routinely collected data are widely used for audit, quality improvement and research. The financial incentives provided by the Quality and Outcome Framework have also boosted computer usage. However, whilst the quantity of diabetes data on diabetes and in many areas management has improved, there has been no exploration of the validity of the underlying data. We carried out this study to define the impact of misclassification or misdiagnosis of diabetes, by exploring the quality of diabetes diagnostic data recorded in routinely collected computer data.

Method

The data examined came from two sources; the Cutting Out Needless Death Using Information Technology (CONDUIT) and Quality Improvement in Chronic Kidney Disease (QICKD). Both of these require the identification of people with diabetes. There are some differences between the populations represented in these databases and also between their populations and a profile of the UK population as a whole. Accordingly different results were found in each but for ease of presentation only the totals covering both CONDUIT and QICKD have been used. The separate analysis can be seen in the full report.

Diagnostic coding issues were reviewed and how non-diagnostic codes could be used to infer the type of diabetes were considered. A number of different methods were used to determine whether they were likely to have diabetes. Three categories of diabetes were used to help audit all the data with all cases being broken down into definite, probable and possible. The data from a wide range of information in a patients’ record was then used to classify cases accordingly. The information used was:

1. A diagnosis of diabetes based on Read codes was broken down into
   a. Definite where a specific code for diabetes, such as C10E for Type 1 and C10F for Type 2, were used with no contradictory codes.
   b. Probable where less specific codes such as that for maturity onset diabetes or more contradictory codes such as those for both Type 1 and Type 2 were present.
   c. Possible where vague high level codes for “diabetes mellitus” were used or where there were multiple contradictions in coding.
2. Therapeutic data broken down into insulin, metformin and other oral anti-diabetic drugs (OADs). The logic used was that everyone with Type 1 should be prescribed insulin, people with Type 2 can be prescribed insulin alone, insulin with an OAD with or without metformin, an OAD with or without metformin, metformin on its’ own or no drug therapy.

3. In the absence of any data on therapy then blood glucose and HbA1c test results were looked for. As blood glucose results are not coded to indicate whether it was fasting or non-fasting sample both 7.1 and 11.1 mmol/l in the past ten years were used to indicate diabetes. For HbA1c the thresholds of HbA1c > 59 mmol/mol (7.5%) and HbA1c >48 mmol/mol (6.5%) were used similarly.

4. Other data such as age, and BMI were also used to sort cases into their most likely category. For Type 1, cut offs at age <35 years were used. (Figures 1 and 2).

Using all the data obtained to sort cases into definite, possible and probable diabetes the following definitions were used to differentiate between proposed changes:

Miscoding is when the wrong computer code is used meaning that it is not possible to determine the type of diabetes precisely.

Misclassification is when someone is incorrectly classified as having a type of diabetes that they don’t have.

Misdiagnosis is when someone is diagnosed with any form of diabetes when they don’t have it at all.

Results

Prevalence

In total just under a million patient records were available to audit; around a quarter of a million in the CONDUIT study; and three-quarters of a million in QICKD (Table 1).

<table>
<thead>
<tr>
<th>Population</th>
<th>All Diabetes</th>
<th>Type 1</th>
<th>Type 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>n</td>
<td>%a</td>
</tr>
<tr>
<td>CONDUIT</td>
<td>220,958</td>
<td>7100</td>
<td>3.2%</td>
</tr>
<tr>
<td>QICKD</td>
<td>760,588</td>
<td>29475</td>
<td>3.9%</td>
</tr>
<tr>
<td>Total/Average</td>
<td>981546</td>
<td>36575</td>
<td>3.7%</td>
</tr>
</tbody>
</table>

a: percentage of records in the database with a diagnosis of diabetes
b: percentage of the recorded diabetes diagnosis with Type 1
c: percentage of the recorded diabetes diagnosis with Type 2
Type 1 Diabetes
The prevalence of Type 1 initially indicated by the two databases was 0.33\% (3,272/981,546) of all cases and 8.9\% (3,272/36,575) of all diabetes cases. The vast majority of these were labelled using a definite Type 1 read code. A major problem with misclassification was found with many people identified as having Type 1 actually having Type 2. In total 1209 people classified as Type 1 were either taking insulin and OAD, OAD alone or no therapy; of these only 129 were plausibly using insulin and metformin meaning that overall 33\% (1,076/3,272) of people classified as having Type 1 did not have it.

Applying the classification methods used to the data reduced the overall prevalence of Type 1 to 0.22\% of all cases and 6.0\% of diabetes cases.

**Figure 1: People with Type 1 diabetes also taking oral anti-diabetic therapy (OAD) – CONDUIT study**

Of 576 “definite” type 1 cases:
- 471 are plausible by virtue of their (insulin alone)
- 8 on insulin + metformin are plausible by virtue of early diagnosis and high BMI

The remaining 97 patients (17\%) seem more likely to have type 2 diabetes
Type 2 Diabetes

The prevalence of Type 2 initially indicated by the two databases was 3.4% (33248/981456) of all cases and 90% (33248/36575) of all diabetes cases. More than 90% of them were classified using definite Type 2 read codes. The analysis revealed 9.3% (3416/33248) cases where a definite code for Type 2 had been allocated but with no treatment for the condition recorded. Of those where no treatment was recorded 53% (1795/3416) had no objective evidence of diabetes. This included the absence of blood glucose testing and diabetes care in the previous ten years.

Applying the classification methods to the data reduced the overall prevalence of Type 2 to 3.3% of all cases and 94% of diabetes cases.

Other forms of Diabetes

The analysis showed very few people coded as having other forms of diabetes. A small number, 14, were identified as having Reaven's (metabolic) syndrome. The number of “others” is almost doubled if we reclassify the young, non-obese apparently with Type 2 into this group.
Summary of misclassification, misdiagnosis and miscoding

Overall misclassification, misdiagnosis and miscoding were present in 14.5% (5,239/36,575) of diabetes cases. Misclassification was the least frequent error accounting for 23% (1,227/5,293) of all errors with the largest number of misclassifications being people having Type 1. Misdiagnosis was more common occurring in 34% (1,808/5,293) of errors with the large numbers of people diagnosed with Type 2 who probably do not have it. The most common error was miscoding accounting for 43% (2,246/5,293) of all errors. Here the principal problem was the use of vague disease codes which do not specify type or contradictory coding within the case record.

Figure 3: Use of an algorithm to identify misclassification, misdiagnosis and miscoding in diabetes: CONDUIT
Figure 4: Use of an algorithm to identify misclassification, misdiagnosis and miscoding in diabetes: QICKD

Total population  
$n = 220,958$

Any diabetes code  
$n = 7100$

Type 1 Definite  
$n = 575$

Type 1 Probable  
$n = 10$

Type 1 Possible  
$n = 13$

Other definite diabetes  
$n = 8$

Reaven's syndrome with euglycaemia  
$n = 5$

Inconsistent treatment  

Undetermined  

Inconsistent treatment

Type 2 Definite  
$n = 5,943$

Type 2 Probable  
$n = 163 + \sim 100$

Type 2 Possible  
$n = 383$

Probably not diabetes  
$n = \sim 300$

Discussion

Principal findings
Currently there are limitations in the way that routinely collected data can be used to confirm or refute a diagnosis of diabetes. Data is often unreliable with date of diagnosis either not shown or being given, after the date therapy started. Treatment and monitoring codes, data about care delivery processes and home testing data are unreliable and rarely add clarity.

As a result it appears that current diagnostic coding practices probably overestimate the prevalence of diabetes overall and also of both Type 1 and Type 2; though there is proportionately greater over-recording of Type 1. The largest problems are the misclassification of people with Type 2 as having Type 1 and diagnosing people as Type 2 when they do not have diabetes at all.
The analysis results have profound implications for diabetes care and treatment. Based on a standard prevalence a practice of 10,000 patients will have between 410-500 people with diabetes. If these results are applied then 60-65 will have some sort of errors and should be flagged for review. There will be 7-16 cases where the wrong sort of diabetes has been diagnosed the majority of which will be incorrectly diagnosed Type 1 when the person really has Type 2. About 21 people will be incorrectly diagnosed with diabetes when they do not have diabetes and there will be between 24-37 miscodings.

**Limitations**

The major limitation is that neither the CONDUIT nor QICKD data may be truly representative of the population as a whole. In CONDUIT the age and gender profile is very different from that nationally. In QICKD the data has been collected from volunteer practices and may not be representative of general practice as a whole.

Also there was no access to the full patient records and conclusions were drawn from a comprehensive but incomplete set of coded data.

**References**


(c) Evidence from the pilot audits
The purpose of this work was to determine whether you could practically apply clinical queries in “live” IT systems and to establish that these queries produced similar figures to those derived from the research databases. It also led to development of the tools described in chapter 7.

Introduction
Coding issues in people with diabetes in routine clinical care have been explored. Gaps in the coding and classification of diabetes as well as for misdiagnosis in data held in primary care records were investigated. The earlier systematic review reported that incomplete coding and classification of diabetes was a known problem¹. Studies of the coding system and surrogate markers for diabetes were also carried out. These studies showed that there are many ways that people with diabetes are represented within current computerised medical records², ³.

Our investigation of routinely collected computerised clinical data used in the CONDUIT and QICKD studies suggested that, of those people with potential errors, between 10% and 25% of people with Type 2 diabetes were incorrectly classified as Type 1; and 5% of those with Type 2 diabetes had no objective evidence of diabetes⁴. We wished to explore whether these cases flagged in our study of miscoding, misclassification and misdiagnosis exist within normal clinical practice.

Method
We developed six MIQUEST (Morbidity Information Query and Export Syntax) queries to test the validity of our principal study findings⁵. MIQUEST is a Department of Health sponsored data extract tool, capable of extracting data from different brands of GP computer system. We ran these queries in “local” mode within individual practices, so that it produced simple data tables of people needing notes review for the practice to process. The MIQUEST query processor and GP computerised record systems cannot carry out the sophisticated data processing carried out in the study of miscoding, misclassification and misdiagnosis paper⁴. Testing also confirmed that the MIQUEST query processor is implemented in subtly different but important ways between the different systems. The MIQUEST manual, for example, says that you can request maximum values; something which is very important in trying to identify if someone has diabetes or not. However, this is not implemented in one of the GP computerised medical record systems we worked with. The implication is that we could not implement queries which used the MIQUEST functionality to the full. Instead we had to produce queries which listed all patients with a condition, treatment or pathology result and then re-order them within a spreadsheet (Microsoft Excel) to identify the high risk patients.

The six queries used are discussed in detail in the section on “Chapter 7(b).

Results
These searches were run in five practices in southeast England with a combined list size of approximately 45,000. The searches flagged 203 out of the approximately 1,600 people with diabetes of whom, 83 had errors. This means only 5% had any errors with 2.2% being misdiagnosed, 2.1% being misdiagnosed and 0.9% being miscoded.
Table 2: Age-sex profile of all cases identified by the audit tools

<table>
<thead>
<tr>
<th>Age bands</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>0-24</td>
<td>5</td>
<td>83.3%</td>
<td>1</td>
</tr>
<tr>
<td>25-49</td>
<td>25</td>
<td>58.1%</td>
<td>18</td>
</tr>
<tr>
<td>50-74</td>
<td>38</td>
<td>35.2%</td>
<td>70</td>
</tr>
<tr>
<td>75</td>
<td>28</td>
<td>60.9%</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>96</td>
<td>47.3%</td>
<td>107</td>
</tr>
</tbody>
</table>

Figure 5: Age-sex distribution of cases identified by the audit

The queries identified more people needing potential changes with Type 2 compared with Type 1 diabetes; simply reflecting the prevalence of these conditions. The numbers and proportion of people identified by each query at the first sort are shown in Table 3. The low numbers identified with query six, reflect technical difficulties in running this query.
Table 3: Number and percentage of people identified by each query

<table>
<thead>
<tr>
<th>Query No</th>
<th>Problem type</th>
<th>Detail</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Misclassification T1DM</td>
<td>No Px Insulin</td>
<td>22</td>
<td>11%</td>
</tr>
<tr>
<td>2</td>
<td>Misclassification T1DM</td>
<td>Px Insulin &amp; OAD</td>
<td>31</td>
<td>15%</td>
</tr>
<tr>
<td>3</td>
<td>Misclassification T1DM</td>
<td>OAD Px before Insulin</td>
<td>6</td>
<td>3%</td>
</tr>
<tr>
<td>4</td>
<td>Misclassification T2DM</td>
<td>T2DM on Insulin from ∆</td>
<td>70</td>
<td>34%</td>
</tr>
<tr>
<td>5</td>
<td>Misdiagnosis T2DM</td>
<td>T2DM but no therapy/abnormal tests</td>
<td>64</td>
<td>32%</td>
</tr>
<tr>
<td>6</td>
<td>Miscoding in diabetes</td>
<td>Use of vague diagnosis codes</td>
<td>10</td>
<td>5%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>203</td>
<td>100%</td>
</tr>
</tbody>
</table>

The types of diabetes identified by the queries also varied. These were additional cases which would not have been included in the QOF disease registers for diabetes. This possibly accounting for why these registers may underestimate the number of cases of diabetes (Table 4).

Table 4: Cases identified by the audit tools with possible coding issues

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1DM</td>
<td>50</td>
<td>24.6</td>
</tr>
<tr>
<td>T2DM</td>
<td>132</td>
<td>65.0</td>
</tr>
<tr>
<td>Unclassified</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Not coded</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td>Vague Diabetes code</td>
<td>18</td>
<td>8.9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>203</td>
<td>100.0</td>
</tr>
</tbody>
</table>

The diagnoses after case review indicated that at least 22 people diagnosed with diabetes did not have the condition (Table 5).

Table 5: Diagnoses after the case review

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Diabetes</td>
<td>22</td>
<td>10.8</td>
</tr>
<tr>
<td>T1DM</td>
<td>32</td>
<td>15.8</td>
</tr>
<tr>
<td>T2DM</td>
<td>126</td>
<td>62.1</td>
</tr>
<tr>
<td>IGT / IFG</td>
<td>7</td>
<td>3.4</td>
</tr>
<tr>
<td>Secondary</td>
<td>13</td>
<td>6.4</td>
</tr>
<tr>
<td>Unclassified</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Gestational</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>203</td>
<td>100.0</td>
</tr>
</tbody>
</table>
There was also a dilemma with another case where the person concerned had no objective record of diabetes within their record other than a single blood glucose reading of 16.4mmol/l. Clearly this result meets the diagnostic criteria for diabetes, but it also raises the question as to whether this might have been a rogue result. Similarly cases were found where the computer record only met the criteria for impaired glucose tolerance or impaired fasting glycaemia (IGT/IFG). However no enquiry beyond the computer record took place and it is possible that further information existed which might have supported this diagnosis.

**Comparing results before and after the audit**

Cross tabulation of the before and after results indicates 40% (20 out of 50) people flagged with Type 1 diabetes actually had Type 2. However five people with Type 2 codes who really had Type 1. Additionally, of the people with a vague diagnosis code for their diabetes (one which could not be readily classified as Type 1 or Type 2) two had Type 1 diabetes and eleven Type 2. Combining these cases with a person with Type 1 with no diagnosis code and the miscoded Type 1 diabetes resulted in an increase in the number of people with Type 1. Five cases of secondary diabetes were found among the cases of Type 2 diabetes. Secondary causes of diabetes are not included as a category within QOF targets, so may be under-represented in GP computerised medical records.

In the four cases where a pre-audit Type 1 diagnosis turned out not to be diabetes: two cases were people who had pancreatic transplants; one a very obese person who had been on insulin at one stage but there was no evidence of diabetes (though a strong family history). The final case just appeared to be a coding error.

**Table 6: Cross-tabulation of pre- and post-audit types of diabetes**

<table>
<thead>
<tr>
<th>Pre-audit</th>
<th>Not DM</th>
<th>T1DM</th>
<th>T2DM</th>
<th>Secondary</th>
<th>IGT/IFG</th>
<th>Un-classified</th>
<th>Gestational</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1DM</td>
<td>4</td>
<td>24</td>
<td>20</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>T2DM</td>
<td>15</td>
<td>5</td>
<td>94</td>
<td>5</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>132</td>
</tr>
<tr>
<td>Un-classified</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Not coded</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Vague DM code</td>
<td>3</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>32</td>
<td>126</td>
<td>7</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>203</td>
</tr>
</tbody>
</table>

Grey shaded cells represent the principal changes as a result of the audit

Overall the effect of the audit was to change the diagnosis of around a quarter of the people identified (table 7).
Table 7: Miscoding misclassification and misdiagnosis of diabetes

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No change</td>
<td>120</td>
<td>59.1</td>
</tr>
<tr>
<td>Miscoded</td>
<td>15</td>
<td>7.4</td>
</tr>
<tr>
<td>Misclassified</td>
<td>33</td>
<td>16.3</td>
</tr>
<tr>
<td>Misdiagnosed</td>
<td>35</td>
<td>17.2</td>
</tr>
<tr>
<td>Total</td>
<td>203</td>
<td>100</td>
</tr>
</tbody>
</table>

Which queries identified cases

The queries which altered the largest number and proportion of cases were inevitably the queries which found cases in Type 2 diabetes (Table 8). However, all the queries identified some patients with problems suggesting that there is heterogeneity in the areas where we need to improve classification of diabetes.

Table 8: Which queries identify cases where there is a change in coding, classification, or diagnosis?

<table>
<thead>
<tr>
<th>Query No</th>
<th>Problem type</th>
<th>Change</th>
<th>No change</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>1</td>
<td>Misclassification T1DM</td>
<td>12</td>
<td>10 1.1%</td>
<td>22 10.8%</td>
</tr>
<tr>
<td>2</td>
<td>Misclassification T1DM</td>
<td>12</td>
<td>19 22.6%</td>
<td>31 15.3%</td>
</tr>
<tr>
<td>3</td>
<td>Misclassification T1DM</td>
<td>3</td>
<td>3 3.6%</td>
<td>6 3.0%</td>
</tr>
<tr>
<td>4</td>
<td>Misclassification T2DM</td>
<td>51</td>
<td>19 22.6%</td>
<td>70 34.5%</td>
</tr>
<tr>
<td>5</td>
<td>Misdiagnosis T2DM</td>
<td>37</td>
<td>27 32.1%</td>
<td>64 31.5%</td>
</tr>
<tr>
<td>6</td>
<td>Miscoding in diabetes</td>
<td>4</td>
<td>6 7.1%</td>
<td>10 4.9%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>119</td>
<td>84 100%</td>
<td>203 100%</td>
</tr>
</tbody>
</table>

Chi-square p<0.001

The ages of people who had a change in their coding was younger; they were also more obese (higher Body Mass Index -BMI); and had a higher glycaated haemoglobin (HbA1c - Table 9). However, these differences were not statistically significant. It is possible that people with diabetes not on QOF registers are less likely to be called in for monitoring and review.
Table 9: Difference in age, BMI and HbA1c between those who had a change and those people with diabetes who had no coding change

<table>
<thead>
<tr>
<th></th>
<th>Change in diagnosis</th>
<th>No change in diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Age</td>
<td>60.1</td>
<td>16.3</td>
</tr>
<tr>
<td>BMI</td>
<td>29.5</td>
<td>6.2</td>
</tr>
<tr>
<td>HbA1c</td>
<td>6.6</td>
<td>2.1</td>
</tr>
</tbody>
</table>

SD = Standard deviation, SEM = Standard Error of the mean, IQR = Inter quartile range

A further comparison of people who were misdiagnosed as having diabetes, compared with the misclassified and miscoded shows that this group, perhaps not surprisingly have a much lower glycated haemoglobin. The people who are poorly classified or coded have a mean HbA1c of 7.39% (SD 1.71; SEM 0.26) compared with an HbA1c of 4.97% (SD 1.49; SEM 0.26) for those we believe don’t have diabetes. This adds some face validity to this claim. These differences are statistically significant (t-test p<0.001).

Manual review of records revealed changes in coding and classification which were not the primary intention of the query. Table 10 shows how the all the queries identified cases where quality improvement was needed, beyond their original scope. For example, the misclassification queries identified a similar number of cases of miscoding of Type 1 diabetes.

Table 10: The extent to which the queries identified the type of cases intended

<table>
<thead>
<tr>
<th>Query No</th>
<th>Problem type intended query identifies</th>
<th>Miscoding n</th>
<th>Miscoding %</th>
<th>Misclassification n</th>
<th>Misclassification %</th>
<th>Misdiagnosis n</th>
<th>Misdiagnosis %</th>
<th>No change n</th>
<th>No change %</th>
<th>Total n</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Misclassification T1DM</td>
<td>12</td>
<td>10.0%</td>
<td>1</td>
<td>6.7%</td>
<td>7</td>
<td>21.2%</td>
<td>2</td>
<td>5.7%</td>
<td>22</td>
<td>10.8%</td>
</tr>
<tr>
<td>2</td>
<td>Misclassification T1DM</td>
<td>12</td>
<td>10.0%</td>
<td>5</td>
<td>33.3%</td>
<td>13</td>
<td>39.4%</td>
<td>1</td>
<td>2.9%</td>
<td>31</td>
<td>15.3%</td>
</tr>
<tr>
<td>3</td>
<td>Misclassification T1DM</td>
<td>3</td>
<td>2.5%</td>
<td>1</td>
<td>6.7%</td>
<td>2</td>
<td>6.1%</td>
<td>0</td>
<td>0.0%</td>
<td>6</td>
<td>3.0%</td>
</tr>
<tr>
<td>4</td>
<td>Misclassification T2DM</td>
<td>51</td>
<td>42.5%</td>
<td>6</td>
<td>40.0%</td>
<td>8</td>
<td>24.2%</td>
<td>5</td>
<td>14.3%</td>
<td>70</td>
<td>34.5%</td>
</tr>
<tr>
<td>5</td>
<td>Misdiagnosis T2DM</td>
<td>38</td>
<td>31.7%</td>
<td>1</td>
<td>6.7%</td>
<td>2</td>
<td>6.1%</td>
<td>23</td>
<td>65.7%</td>
<td>64</td>
<td>31.5%</td>
</tr>
<tr>
<td>6</td>
<td>Miscoding in diabetes</td>
<td>4</td>
<td>3.3%</td>
<td>1</td>
<td>6.7%</td>
<td>1</td>
<td>3.0%</td>
<td>4</td>
<td>11.4%</td>
<td>10</td>
<td>4.9%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>120</td>
<td>100%</td>
<td>15</td>
<td>100%</td>
<td>33</td>
<td>100%</td>
<td>35</td>
<td>100%</td>
<td>203</td>
<td>100%</td>
</tr>
</tbody>
</table>

Chi-square p<0.001
Conclusion
This pilot audit shows that there is scope to improve diabetes data quality and that queries of this type may help achieve such a rise in data quality. The changes found were heterogeneous and often required judgement. There is going to be no easy single fix of data quality issues – though this comment needs to be contextualised within a context where UK primary care records are among the best in the world.

Although we did not demonstrate a statistical difference it is plausible that there may be a link through not being on the QOF register and sub-optimal care. The general practice computerised medical record system vendors have developed the capability of their systems to support QOF – usually these system improvements include prompts and structured audit tools. It would almost be surprising if these tools did not demonstrate some benefit.

QOF may distort coding. Payment in diabetes for being on the QOF register is proportional to the number of cases on the register, modulated by the number of quality points and a deprivation factor. There is currently no incentive to code secondary or gestational diabetes correctly; indeed correct coding may undermine income! Discussions are needed to explore whether the algorithm in this document should become part of the QOF recognised denominator population for the diabetes indicator.

An effective diabetes data quality audit tool is going to need to have the capacity to process data outside the clinical record – as several factors need to be weighed up in deciding the type of diabetes. In particular data at the onset of the condition is often lacking. Meanwhile a toolkit of queries, which produce output requiring some further basic ordering, has a role in improving quality. Practices and localities are invited to consider running these or other improved versions of these searches to improve data quality in people with diabetes.

Diabetes clinics in secondary care – may also have a role in helping to more clearly flag the type of diabetes.

Part of every practice’s preparation for its annual diabetes review should be a review of the data quality of its people with diabetes – including whether or not they are correctly coded. Details of how to access and download our pilot queries are contained in Chapter 7(b) Support Tools.

References


6. Improving Diagnosis

This section has been edited by Professor Andrew Hattersley, Professor of Molecular Medicine, Peninsula Medical School, University of Exeter.

(a) Developing guidelines for classification

Understanding potential reasons for problems with classification

There are many potential reasons why problems with classification of diabetes occur. Some errors are simple mistaken entries in the clinical records when the clinical team and patient both know the correct classification. These are unlikely to have an impact on patient care. However there are also errors which result from a lack of information or a lack of understanding by the healthcare professional. The most widespread misunderstanding is to change a diagnosis from Type 2 diabetes to Type 1 when insulin treatment is started. This, potentially, could have a considerable impact on patient care as the guidelines for insulin use in Type 2 are very different from those in Type 1. Finally there are patients in whom it is difficult to make a correct classification, particularly at diagnosis. The rapid increase in obesity has resulted in difficulties in differentiating Type 1 from Type 2 diabetes.

The classical criteria of age of diagnosis, severity of hypoglycaemia, presence of ketoacidosis and BMI provide far less separation between the two main aetiological subgroups than they did when obesity was less prevalent in the population. This is probably best indicated by a fictitious example of Mr 33. He presented with glucose of 33mols/l at the age of 33 with a BMI of 33kg/m². Such a patient may have Type 1 diabetes and will rapidly become totally insulin deficient. Alternatively he may have Type 2 and, with early dietary measures and oral agents such as metformin, achieve excellent blood glucose control. The key thing here is to recognise when there is uncertainty in classifying the type of diabetes people have and to separate the immediate clinical management of the patient and the subsequent classification of the diabetes. In cases that are hard to classify the response to initial treatment, as well as aetiological and other investigations, can subsequently help to make an accurate classification. Often the hardest time to diagnose the type of diabetes is at diagnosis.

A core problem is the lack of national or international, clinically based guidelines for the classification of diabetes. Documents on classification produced by the American Diabetes Association¹ and the World Health Organisation² have opted to use aetiology of diabetes as the most rational approach to classification. These documents do not give precise clinical definitions predominantly because of the difficulty in providing accurate criteria. For example Type 1 diabetes can occur at any age, hence any proposed cut of age will not be correct in all cases. In addition, the fact that there are not clear gold standards based on investigations, particularly for the definition of Type 2 diabetes, makes it very hard to assess the sensitivity and specificity of any clinical criteria used. Probably the clearest criteria for Type 1 diabetes is that ultimately, in almost all cases, absolute insulin deficiency¹,² is found although the time taken for this to occur is very variable. However this cannot be used soon after diagnosis and is rarely tested in clinical practice. One solution to an aetiological classification might be the increased use of aetiological investigations. In the UK the measurement of pancreatic auto antibodies at diagnosis in patients with probable Type 1 diabetes is far from universal and both false positives and false negative results occur³⁴⁵. One situation where aetiological tests clearly do work is in the definition of genetic subtypes where correct classification often changes the optimal treatment⁶. However their rarity means that this form of testing can only be used where the diagnosis is suspected and many clinicians do not have adequate exposure and training to recognise these subgroups.
Finally there is a misunderstanding with patients who do not meet the criteria for diabetes being included in some schemes for the classification of diabetes. This includes the subgroups at increased risk of diabetes often with borderline hyperglycaemia such as impaired glucose tolerance, impaired fasting glucose and gestational diabetes. In gestational diabetes if diabetes persists outside pregnancy this should then be classified according to the subtype of diabetes or if it resolves then the patient is in an at risk group of developing Type 2 diabetes in the future.

**Why does it matter?**

It is important that patients are classified correctly. It is critical for the appropriate management of the patient. As the guidelines for managing Type 1 and Type 2 diabetes differ greatly this means the incorrect classification will lead to an inappropriate management plan. Other examples are that the most common subgroups of MODY and neonatal diabetes are frequently misdiagnosed as Type 1. This results in patients, many of them children, being treated with insulin injections when they often achieve better glycaemic control with sulphonylurea tablets and it is important to identify these. Earlier in this document the importance of a correct classification for patients and also the difficulty and uncertainty that results when the classification varies with different health care staff has been explained. Finally accurate classification is important for the population estimates of present and future healthcare needs and for defining research priorities.

**Finding a solution**

The solution to the problems of diabetes classification cannot be realistically dealt with by all patients receiving long-term review by diabetes experts or widespread use of sophisticated tests. While there is the crucial role of the diabetes expert in cases where classification is uncertain, for many patients classification could be helped by the use of clinical guidelines. The aim was to develop simple clinical guidelines that could be used in general practice with material that was routinely available in practice records. It is not intended to replace expert opinion or aetiological testing but rather to give a starting point for healthcare staff. It aims to be pragmatic and simple. Key components were to separate diabetes from non diabetes hyperglycaemia, to include broad categories for genetic and secondary forms of diabetes and the inclusion of an undefined category. The guidelines are summarised in figure 6.

It is appreciated that patients may change categories as more is understood about their diabetes. For instance an obese patient with Type 2 diabetes who is subsequently found to have Cushing’s disease will change from Type 2 to a diagnosis of “other”.

**Limitations of guidelines**

When writing guidelines it is easy to understand why other expert bodies have chosen not to produce simple clinical criteria. Inevitably, for any suggestion made it will be easy to identify cases that do not conform to the proposed classification. The classification used is primarily based on age of diagnosis and treatment, which means there are other clinical factors such as the degree of obesity which will help classification; however such factors have been less discriminatory than age of diagnosis and treatment. It is important to stress that it is not thought that such guidelines should or could replace expert assessment including aetiological investigations.

The broad categories used, simplify the classification but also means in some cases they will not be precise enough to define a patient’s diabetes. Patients who have a pancreas removed for pancreatitis or have diabetes because of acromegaly would both be defined as “Other” but would require very different treatments.

The aim of this is to provide a simple pragmatic solution that can be used in primary care. Ultimately its worth will be shown by its utility in clinical practice and if its use results in a reduction in the degree of misclassification. It is crucial that, as with any other guideline, they are reviewed and modified over time.
### References


7. Improving existing patient records

This section has been edited by Professor Simon de Lusignan, Professor of Primary Care & Clinical Informatics, University of Surrey.

(a) Good coding practice

“...we think in narratives but computers think in codes!”

Introduction

The changing role of primary care requires high quality clinical records:

The role of primary care and the record systems needed to support patient care has changed over time (Figure 7). The initial role of primary care was as a reactive and accessible service; single-handed or small practices provided local care and medical records were often a brief aide memoire for the physician. Primary care has evolved from a universally accessible service, as defined by the World Health Organisations (WHO) Alma Ata declaration of 1978. The first stage in evolution was the requirement for specially trained professionals in general practice. Subsequently general practice has evolved to being a provider of comprehensive services; and most recently towards being an integrator addressing the majority of healthcare needs in the context of family and community, including being accountable for how it provide services and uses resources.

The expanding role of primary care and its need to address quality has been codified in the WHO more recent Almaty declaration: Primary Care Now More Than Ever.

Figure 7: Changes in definition and function of primary care

- Universally accessible
  - (Alma Ata, WHO 1978)

- Level of care + training
  - (Fry 1980)

- Set of attributes
  - First contact, Longitudinality,
  - Comprehensiveness
  - (Starfield 1992)

  - Integrated, accessible, accountable,
  - Addressing majority of healthcare needs
  - Context of family + community
  - (IOM 1996)

- About partnerships
  - Expanding problems dealt with + quality
  - (Almaty – Now more than ever, WHO 2008)
Primary care records have needed to evolve with this change. There is international consensus that high quality records are needed to underpin the development of health systems. The medical record has evolved to meet the needs of primary care.

The key evolutionary steps in the development of the medical record have been:

- The creation of the patient centred record (one record per patient). This took place in the 19th century; the first place to do this systematically was the Mayo Clinic in the United States
- Registration systems, where the creation of NHS numbers allowed records to follow patients as they moved around the health system
- Special records with special data entry forms for the record summary, immunisation and management of long term conditions were created, initially on paper, but later on computer
- Problem orientation of records where records linked problems with treatment – usually diagnosis with therapy
- Sharing of records across the primary health care team where records are shared between doctors, nurses and other members of the primary health care team involved in the care of patients. More recently computerised records provide the opportunity for records to be shared much more widely
- Patient access to records where people can have access to their own medical record as a way of improving accuracy and supporting patient empowerment.

For more integrated care to take place, clinical data recorded in one place may need to be used in another. The clinical record no longer just needs to be understood by the clinician who made the record or their immediate colleague who understands the context within which the record was made. Instead the record is potentially used at greater distance. This distance can be considered in three dimensions: distance in person, place and time (Figure 8).

**Figure 8: Evolution of the medical record – over time records are used at greater distances from their point of origin. Distance has three dimension: physical distance, time, and person**
Only computerisation and standardisation of the record makes it feasible for information to be shared across the healthcare system. Good clinical coding is a key underpinning of a searchable and sharable computerised record.

**Computerised medical records – primary and secondary use of data**

Computerised medical records facilitate audit and quality improvement on a scale that was not feasible with paper records. However to do this they must contain good quality data and be appropriately linked into their wider health system.

They can be linked to information and decision support; to systems for flagging drug interactions; and can be linked to other systems such as outpatients booking and Choose and Book; pathology laboratories for results and can be searched to identify poorly managed cases.

Data quality is critical, although there currently is no single accepted definition. The preferred definition of data quality is that developed by the Primary Care Informatics Working Group of the European Federation for Medical Informatics (EFMI)8 which is that data should be “Fit for purpose.” This broad definition encompasses the wide range of legitimate primary and secondary uses of clinical data. The primary use of clinical data is to directly support patient care. Secondary use of data, usually in anonymised form provides practices and health service managers with prescribing and referral data; information about the quality of disease management; data for other types of health service management; quality improvement initiatives; and research9 (Figure 9).

**Figure 9: Examples of secondary uses of primary care data**
Types of clinical data in computerised records: “Coded” and “Free text”

Medical records contain administrative and clinical data. The clinical data in computerised medical records is of two sorts:

(1) The coded part uses a standardised list of terms to ensure that all the different ways a disease, or other clinical concept, might be represented are coded using the same code or group of codes\(^1\). For example, the Read code for diabetes is “C10”; the code for Type 1 diabetes is “C10E” and the code for Type 2 is “C10F”.

(2) The free text, or narrative, record is typed into the appropriate part of the medical record. The history is generally a narrative record as is the clinician’s management plan. Examination findings are usually a combination of codes and narrative text.

Figure 10: Components / types of data in a computerised medical record system

Coding data is critical

Searching and monitoring quality on a GP computer system is dependent upon recorded coded data. Computers can readily process coded data, but not free text. Although the science of natural language processing (NLP) is improving all the time it is not yet ready to obviate the need to code data. For the foreseeable future only the “coded” part of the record is readily searchable.

However, whilst we need to ensure that data are coded, we also need to recognise that coding data is not always easy\(^1\). Coding tends to have a biomedical focus, whereas many primary care consultations have a more psychological and social components. Some patients react badly to some diagnostic labels, for example they might not want to be labelled as depressed or to have asthma.
Creating a properly coded record is a professional duty not an option

The General Medical Council’s duties of a doctor include that all registered medical practitioners should make a proper medical record at the time of the clinical consultation\textsuperscript{12} and other professional groups make similar stipulations. Specific recommendations about computerised records can be found in RCGP\textsuperscript{13} and RCGP/DH guidance\textsuperscript{14} and there are also recommendations from the summary care record programme\textsuperscript{15}. In the NHS in England there has been a centrally funded programme to bring records up to a standard to enable them to contribute summary health data to a shared summary care record which is available throughout the NHS should a patient ever fall ill.

Components of a clinical record

An adequate computerised medical record should include a problem title for every consultation, which is usually coded; and sufficient detail about the other elements of the consultation that the record can be used by another practitioner to provide continuity of care. Complete demographic data including ethnicity; preventive, screening, risk factors, therapy and diagnosis recording can all contribute to monitoring the quality of care. Data entry forms, sometimes called templates, facilitate computer data entry.

Continuous improvement of the clinical record

Computerised records need continuous improvement and maintenance. There are several areas where record content can be improved:

(a) Diagnosis / Problem title recording – every patient should have a problem title for every consultation.
   i. Where possible record a disease code, in the Read coding system where disease codes start with a letter. Only change the problem title if it adds something to the care of the patient. Thus C100 – “Diabetes with no mention of complications” is usefully modified to C10EA “Type 1 diabetes with no complications” (assuming they have Type 1).
   ii. If you can’t make a diagnosis use a symptom code. Ideally one from the “R” chapter such as R0814 Polyuria.
   iii. Only use other codes, as the problem title, if it is impossible to use a diagnosis or symptom code.

(b) Avoiding recording of multiple similar codes for the same diagnosis; most computerised systems allow you to group together or unify patients with multiple codes to a single diagnostic code.

(c) Duplicate recording of the same event.

(d) Linking diagnosis with the relevant prescription.

(e) Ensure capturing clinical data recorded outside the clinical record, which is not normally coded (e.g. in letters or other reports).

(f) Looking at ways of increasing the amount of data coded, without increasing the workload on clinicians.

Good coding practice and diabetes

There are some specific issues in good coding practice in diabetes:

(a) Getting the diagnosis right: It is often quite challenging to get the diagnosis right with diabetes. Under some circumstances it can be difficult to differentiate Type 1 from Type 2, and, also to be certain that someone actually has diabetes at all.
(b) Evolving classification and diagnostic criteria for diabetes: The classification of diabetes has changed over time, as has its diagnostic criteria. Some years ago practitioners described both juvenile type, and maturity onset; additionally they talked about non-insulin dependent diabetes (NIDM) and Insulin dependent (IDM). Although no longer used, these old diagnostic codes, unfortunately, have to remain in the coding system. If they were removed items coded using these terms will become “orphans” - clinical data without any links in the coding system. This means that on-going maintenance and care of redundant codes is required.

(c) Interpreting clinical data – especially blood glucose results: The origin of results may be unclear: Who did the test? Was the test carried out when the patient was fasted? Was the test part of a glucose tolerance test? Was the patient pregnant? Plasma glucose results carried out in a pathology lab are likely to be reliable, whereas those carried out using stick tests may be less accurate. When computerised records or results are searched the origin of the data may not be taken into account – unless fasting and non-fasting specimens are scrupulously separately coded.

(d) Distortion of coding by pay-for-performance: QOF has more indicators for diabetes than any other clinical area; with the aim of incentivising improvements in chronic disease management. Only people with Type 1, coded with C10E hierarchy codes, or Type 2, coded within the C10F group of codes, are included in the pay for performance disease register. The coding of people with other diabetes codes from the C10 family means they will not be included in this register and may not benefit from its prompts and incentives. It is also possible that in some practices everyone with diabetes was given a Type 1 or Type 2 diagnostic code and this may have inevitably led to errors.

Among the duties of a doctor are keeping high quality records and accurately coded data are the key components of a high quality record. Our record quality needs to match the needs of our patients in the current health system – where sharing of records and coordinating if not integrating care will become the norm. In the “First Law” of Informatics van der Lei wrote that data can only be used for the purpose for which it was recorded. A sustainable position where records are only used within a small organisation only involved in the care of patients. However, within a modern health system where we are required to assess health needs and ensure equity and wish to commission care for communities we need to share and make secondary use of data. A “Second Law” of informatics states so long as you fully understand the context within which data are recorded you can ignore the First Law (Figure 11).

Figure 11: The first and second laws of informatics

- The first law of informatics: (van der Lei 1991)
  
  Data shall be used only for the purpose for which they were collected... ...if no purpose was defined prior to collection... ...the data should not be used

- The second law of informatics: (lusignan & Mimnagh 2006)
  
  If you fully understand the context of data recording you can ignore the first law!

There are particular issues in coding data in diabetes: Changes and difficulties in classifying some people and the lack of contextual recording around plasma glucose records. However, keeping records to high standards requires ongoing maintenance. That said UK primary care computer records are among the highest quality in the world and our routinely collected data are widely used for research.
References
macauley-of-mice-and-real-people/

2. Schade CP, Sullivan FM, de Lusignan S, Madeley J. e-Prescribing, efficiency, quality: lessons from the

3. de Lusignan S, Chan T. The development of primary care information technology in the United kingdom. J
Ambul Care Manage. 2008 Jul-Sep;31(3):201-10.


Press; 1996

7. de Lusignan S, Teasdale S, Little D, Zapp J, Zuckerman A, Bates DW, Steele A. Comprehensive computerised
primary care records are an essential component of any national health information strategy: report from an

8. de Lusignan S. The optimum granularity for coding diagnostic data in primary care: report of a workshop of the

9. 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. URL:
http://fampra.oxfordjournals.org/cgi/reprint/23/2/253

10. de Lusignan S. Codes, classifications, terminologies and nomenclatures: definition, development and application

11. de Lusignan S, Wells SE, Hague NJ, Thiru K. Managers see the problems associated with coding clinical data as a

March 2009.

13. RCGP and GPC (GP Committee of British Medical Association). Good Medical Practice for General Practitioners.

14. The Joint General Practice Information Technology Committee of the General Practitioners Committee and
RCGP. Good practice guidelines for general practice electronic patient records (version 3.1). London;
Department of Health (Ref: 5098), 2005. URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/
@dhw@en/documents/digitalasset/dh_4116707.pdf

15. Connecting for Health. Summary Care Records Data Quality – Interim Requirements Statement. URL:

16. de Lusignan S. Flagging fasting plasma glucose specimens: time to routinely label the context in which
pathology specimens are recorded. Inform Prim Care 2009;17(2):63-4.

17. de Lusignan S, Mimnagh C. Breaking the first law of informatics: the Quality and Outcomes Framework (QOF)

18. van der Lei J. Use and abuse of computer-stored medical records. Methods of Information in Medicine
(b) Supporting Tools for existing patients with diabetes

Introduction
This section takes the first steps towards creating a toolkit to identify people with possible miscoding, misclassification or misdiagnosis of diabetes. It brings together all the learning from this project. The aim of this work is to create high quality records which can underpin improved quality of care for people with diabetes. This chapter describes:
• The first step towards creating a toolkit to flag cases for review
• The queries used to extract data for the pilot audit providing the basis for this toolkit
• The pilot toolkit we have created to use with the EMIS LV – brand of general practice computerised medical record system.

What was known already?
The use of Boolean logic to identify cases which may be appropriately or inappropriately coded is well established in informatics. A “data quality probe” is a two-by-two matrix to identify whether a code is correct or incorrect. Often data quality probes will look at the relationship between the presence or absence of a medicine compared with the recording or not of a disease code (e.g. Prescription of insulin vs. Type 1 Diabetes); they can also look for the absence of a condition. The Care and Health Analysis in Real Time (CHART) tool adopts this principle and is widely used to improve data quality and raise records to the standard where they are fit to be shared across the NHS. This approach is useful, though limited in the complexity of what it can explore. For example, where it is unclear whether a person has Type 1 or Type 2 diabetes, other factors including age, ethnic origin and BMI have been used to differentiate types of diabetes.

Use of downloadable quality improvement tools as for quality improvement
Downloadable quality improvement tools have been provided previously. These have been widely used, though practices have often required additional expertise, from people with MIQUEST and data processing expertise, to run them effectively. Downloadable MIQUEST queries have been provided for the Improving Access to Psychological Therapies (IAPT) programme and also for calculating renal function – estimating glomerular filtration rate (eGFR) before this test was widely available at laboratories.

Method
Overview
The strengths and limitations of medical record systems are well known. We developed logical models to develop the queries utilising the diagnostic rules within the algorithm for the diagnosis of diabetes; modulated by the strengths and weaknesses of medical record systems and the data extraction tool MIQUEST. A number of features of an individual case of diabetes are used to determine the type. This is a relatively complex process to classify a patient’s diabetes. For example: A person with Type 1 diabetes will usually have been prescribed insulin within 6 months of diagnosis if under 35, and from diagnosis if older; People with Type 2 diabetes are unlikely to have started insulin with 6 months of diagnosis if under 35 years, and to have some breaks in therapy.
Read code list
A list of clinical codes, in addition to age and gender, needed to use to differentiate types of diabetes using the algorithm were identified. The first generation of tools are using 5-byte Version 2 Read codes, the codes used by most computerised medical record systems, though for these queries to be used across all systems the equivalent for Clinical Terms version 3 will need to be developed; a version used by a minority of system vendors.

Fasting glucose blood tests
Fasting blood sugar 44T2
Fasting blood glucose 44TKI

Random glucose blood tests
Random blood sugar 44T1
Laboratory blood sugar 44T5
Plasma glucose 44TA
Blood glucose level 44TJ

Glycated haemoglobin – HBA1c
42W5 HbA1c level (IFCC aligned)
42W4 HbA1c level (DCCT aligned)

There are other results within the 44T. code set which might also be associated with an HbA1c value. As these are older codes these are likely to have DCCT the old “%” units assigned to them.
44TB Haemoglobin A1c level
44TC Haemoglobin A1 level
44TL Total glycosylated haemoglobin level

Oral Anti-Diabetic drugs (N.B. EMIS LV uses a different drug dictionary):
The hierarchy is the endocrine drugs:
f.... ENDOCRINE DRUGS

This hierarchy includes:
f1... SHORT-ACTING INSULIN PREPARATIONS
f2... MEDIUM/LONG-ACTING INSULINS
f3... SULPHONYLUREAS
f4... BIGUANIDES (This is metformin)
f5... GUAR
f7... HYPOGLYCAEMIA TREATMENT (Rarely prescribed – but f72..
GLUCAGON is more common for self administration)
ft... OTHER DRUGS USED IN DIABETES (This is the new gliatazones, some are combined with metformin but this does not matter)
fw... SHORT WITH INTERMEDIATE-ACTING INSULINS
Incomplete data – moving practice & attending diabetes clinic

Incomplete data can exist for a number of reasons. However, we found two major problems: firstly in recently registered patients or patients who have moved practice since diagnoses as having diabetes and breaks in data while patients are seen in the hospital clinic.

Data may be incomplete, especially when a patient moves practice – partly because it is time consuming to re-summarise records and also because it is not practicable to load up all of a patients prescription data. Although electronic transfer of records is now possible this will only benefit patients who transferred practice in the last couple of years. When a patient with diabetes moved practice prior to electronic record transfer being available it is highly likely that their diabetes diagnosis would be added to their summary by their new practice. However, it is extremely unlikely that all their prescribing data would be transferred. Hence, the data of their diagnosis of diabetes would be in their new record, but not the date they started therapy. The second major problem with diabetes data is that in many practices initiation of insulin is carried out in hospital diabetes clinics, who also prescribe insulin. This means that there may not be an insulin prescription in the GP record at the time of diagnosis. Similarly breaks in insulin prescription, where records are continuous, may be due to care being provided in hospital.

Inconsistent implementation of MIQUEST in different brands of computerised record system

The lack of consistent functionality limited the ability to produce highly focussed outputs which would not require further sorting and processing post extraction. MIQUEST is not identically implemented by the different GP vendors. MIQUEST is written in a database language called health query language (HQL). Some of the most useful commands in HQL are the queries that request maximum or minimum values. This command is effective in some brands but not others.

MIQUEST also outputs dates in standard database format: yyyymmdd.; for example 28th June 2010 would appear 20100628. (This format is used because it allows dates to be sorted numerically).

Pilot query development

Six MIQUEST queries to identify cases with coding problems for each of the six main General Practice IT systems were developed. The details of the method are set out in the section on the pilot audits (Section 5.c.

The six queries developed are as follows:

1. Misclassification of type 1 diabetes

Logic from algorithm: Everyone with Type 1 diabetes should be prescribed insulin. People with Type 1 diabetes require insulin for survival.

Query aims to identify: People not on insulin and people who have not received an insulin prescription for over 6 months with a Type 1 diabetes code.

Possible exceptions: People attending and being prescribed their insulin from specialist clinics.

Query to list: Subset of Type 1 Diabetes (C10E%), and list date, rubric, value 1, text (and code) for latest insulin Px. The output file is viewed to see whether there are diagnostic codes for people without insulin being prescribed.
2. Misclassification of Type 1 diabetes

Logic from algorithm: People on insulin and an oral anti-diabetic medicine are likely to have Type 2 diabetes.

Query aims to identify: People with a diagnostic code for Type 1 diabetes who are concurrently prescribed oral anti-diabetic medicine (OAD).

Possible exceptions: People with multiple codes; a small number of people with Type 1 diabetes may legitimately be prescribed metformin for weight loss.

Query to list: Subset of people with Type 1 Diabetes, then subset on insulin who are also taking an OAD, other than metformin. The query will list people taking OADs and insulin – based on their latest prescription. This query also detects people with a Type 1 diabetes code who are just taking an OAD.

Possible exceptions: In some people it is genuinely difficult to tell the difference between Type 1 and Type 2. This may happen when Type 1 diabetes starts in adult life in someone who from their morphology is expected to have Type 2.

3. Misclassification of Type 1 diabetes

Logic: People who started OADs before insulin are unlikely to have Type 1 diabetes.

Query to list: People with Type 1 diabetes will have started insulin within 6 months of diagnosis. The output file has separate columns for first insulin and first prescription of an OAD. The data need to be viewed and cases highlighted where the OAD prescription date is more than six months before the insulin prescription date.

Possible exceptions: As query 2.

4. Misclassification of Type 2 diabetes

Logic: People who have required insulin from within 6 months of diagnosis are unlikely to have Type 2 diabetes. People with Type 1 diabetes will have continuous therapy with insulin.

Query to list: First and latest insulin prescription, and the date of first OAD for people with Type 2 diabetes. The query has a similar format to query 3, but is a much more extensive list. The output file has to be sorted into those who have insulin before an OAD; followed by examining the diagnosis date for this group of patients. The query also identifies people with Type 2 diabetes who have never been prescribed an OAD in the practice.

5. Misdiagnosis of Type 2 diabetes

Logic: People with Type 2 diabetes will have been prescribed therapy for diabetes (we will exclude metformin); or had abnormal tests – plasma glucose and HbA1c; or be prescribed testing kits for blood or urine.

Query to list: This query was intended to list the highest value for HbA1c and plasma glucose for people with Type 2 diabetes, not prescribed an OAD (excluding metformin) and not prescribed insulin.

Possible exceptions: The maximum command does not work within the MIQUEST implementation on all brands of computer system. This means that the query has had to be modified to display the earliest and latest readings. These then need to be sorted manually in a spread sheet – looking for the highest values.

Fasting plasma glucose tests are inconsistently coded. We therefore suggest a conservative sort using a threshold of <7.0mmol/l representing “no evidence” of diabetes. As it is impossible to tell for most specimens
if they are extracted fasting or not. This sort should be followed by a sort of glycated haemoglobin, concentrating only on the group of people with a plasma glucose of <7mmol/l.

The recommended search criteria is HbA1c <6.5% or 48mmol/mol for Diabetes Control and Complications Trial (DCCT) and International Federation of Clinical Chemists (IFCC) aligned assays respectively.

People with glucose <7.0mmol and HbA1c <6.5% or <48mmol/mol are categorised as unlikely to have diabetes.

6. Miscoding in diabetes

Logic: Some people with diabetes are miscoded with vague codes which imply diabetes but do not list the type.

Query aim: To list people with diabetes who only have vague diagnostic codes (C10 variants – but no use of C10E or C10F) and also to list those with both C10E and C10F (i.e. simultaneously coded for Type 1 and Type 2 diabetes).

Query to list: Query to list: Subset of people labelled with other C10 (diabetes) code who are not C10E and also not C10F! This query produces a large output file. This file should list all the patients who have C10 (diabetes) codes and not C10E (Type 1 diabetes) and not C10F (Type 2 diabetes). However, in theory the logic of the query should have allowed those with both codes to be listed.

Possible exceptions: The query produces a large output which needs to have multiple sorts. Again we recommend sorting by C10E first and then C10F to identify people with contradictory codes (C10E and C10F).

The next group to identify are those with no C10E or C10F code. These people have diabetes but their diagnosis is recorded using codes not included in the QOF disease register, so won’t be picked up by the QOF recall systems.

Results

All of the queries identified cases; and there was inevitable overlap in the results from between queries 1 to 3. However, each patient is only listed once and the practitioners carrying out the audits were asked to only assign each case to one query. Each patient only had one audit form – so there was no chance of double counting. The audits were carried out by GPs or practice nurses who were involved in supporting their practice diabetes clinic. The participants in the audit required help to sort the query output in a spreadsheet, Microsoft Excel, and highlight the cases they should look at. None felt they would have the confidence to sort these data without that expert assistance.

Table 11: Number and percentage of people identified by each query

<table>
<thead>
<tr>
<th>Query No</th>
<th>Problem type</th>
<th>Detail</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Misclassification</td>
<td>T1DM</td>
<td>22</td>
<td>10.8</td>
</tr>
<tr>
<td>2</td>
<td>Misclassification</td>
<td>T1DM</td>
<td>31</td>
<td>15.3</td>
</tr>
<tr>
<td>3</td>
<td>Misclassification</td>
<td>T1DM</td>
<td>6</td>
<td>3.0</td>
</tr>
<tr>
<td>4</td>
<td>Misclassification</td>
<td>T2DM</td>
<td>70</td>
<td>34.5</td>
</tr>
<tr>
<td>5</td>
<td>Misdiagnosis</td>
<td>T2DM</td>
<td>64</td>
<td>31.5</td>
</tr>
<tr>
<td>6</td>
<td>Miscoding in diabetes</td>
<td>Use of vague diagnosis codes</td>
<td>10</td>
<td>4.9</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>203</td>
<td>100</td>
</tr>
</tbody>
</table>


All the queries identified cases from their output (Table 11). In around 40% of cases identified by the queries and looked at within the audit the patient’s disease code was changed. Within this 40% (n=83) of cases 18% (15/83) were miscoding, 40% (33/83) misclassification and 42% (35/83) misdiagnosis. Table 12 shows which queries identified changes.

Table 12: Which queries identify cases where there is a change

<table>
<thead>
<tr>
<th>Query No</th>
<th>Problem type</th>
<th>Change</th>
<th>No change</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>1</td>
<td>Misclassification T1DM</td>
<td>12</td>
<td>10.1%</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Misclassification T1DM</td>
<td>12</td>
<td>10.1%</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>Misclassification T1DM</td>
<td>3</td>
<td>2.5%</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Misclassification T2DM</td>
<td>51</td>
<td>42.9%</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>Misdiagnosis T2DM</td>
<td>37</td>
<td>31.1%</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>Miscoding in diabetes</td>
<td>4</td>
<td>3.4%</td>
<td>6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>119</td>
<td>100%</td>
<td>84</td>
</tr>
</tbody>
</table>

The queries identified people with miscoding, misclassification and misdiagnosis. All but one of the queries identified both cases of misclassification and miscoding. Misdiagnosis was much rarer.

Pilot toolkit

We have developed a toolkit for the each of the six main GP IT systems. These are

- EMIS LV
- EMIS PCS (LAN+Enterprise edition)
- INPS Vision
- iSoft Premier/Synergy
- Advanced Crosscare
- TPP Systm One

The toolkit contains three components:

1. Comprehensive guide to running the queries
2. An Excel spreadsheet that they can be incorporated into
3. Macros which automatically sort the query output and highlight the cases for review.
4. Links to useful further reading about the research which underpins this audit.

The toolkit is downloadable from http://www.clininf.eu/cod a screen shot of the webpage is shown in Figure 12.
The queries, workbook and spreadsheet are freely available for non-commercial use in the NHS. However, their origin should be acknowledged whether used in whole or part.

**Discussion**

The queries were designed to provide a toolkit for practitioners to use to check for miscoding, misclassification and misdiagnosis of diabetes. They were also designed to test whether the patients identified as miscoded, misclassified and misdiagnosed using coded data held good when the cases identified were checked against their entire medical record.

The queries as currently designed would be challenging for many practitioners to use. All five practices who participated in the audit required expert assistance to sort the output and flag the people that needed investigating. They also did not directly test the logic used in the miscoding, misclassification and misdiagnosis paper; instead they followed more closely to the logic set out in the algorithm in chapter 6. However, that said, over a quarter of the patients flagged within the audit queries had clinically significant changes in their coding, classification or diagnosis.

The practices who participated in the pilot audit may not be representative, in that they may have higher standards of diabetes management and hence the audit may identify fewer cases. The practices that tested these queries are all large; the area has a higher number of working age people, with less retired and, being more affluent, a lower prevalence of cardiovascular disease and diabetes than nationally. They are also all training and teaching practices in an area with a stable practice population – generally practices in this locality turnover less than 5% of their registered list each year. The practices also all have practice-based diabetes
clinics which are supported through an educational programme run by the local diabetes lead consultant supported by a team of diabetes specialist nurses. The practices have all had support to send their lead clinicians on the Warwick primary care diabetes course. All these factors suggest that diabetes care is likely to be good, and that if these audits identify clinically useful cases in these practices they should do the same anywhere.

Please send any comments, questions thoughts about how to improve these queries to: Simon de Lusignan at s.lusignan@surrey.ac.uk

References
2. PRIMIS+. CHART (Care and Health Analysis in Real Time). University of Nottingham. URL: http://www.primis.nhs.uk/index.php/services/information-services/chart
5. St. George’s – University of London. Primary Care electronic Library – eGFR Calculator. URL: http://www.pcel.info/gfr/
10. de Lusignan S. Flagging fasting plasma glucose specimens: time to routinely label the context in which pathology specimens are recorded. Inform Prim Care 2009;17(2):63-4.
Appendix A  Diabetes Classification Guidelines

Practical Classification Guidelines

Diabetes

Type 1
Diagnosis < 35 years* AND continual insulin treatment within 6/12 months of diagnosis OR Diagnosis ≥ 35 years* AND continual insulin treatment from diagnosis

Type 2
Diagnosis < 35 years* AND not on continual insulin treatment within 6/12 months of diagnosis OR Diagnosis ≥ 35 years* AND not on continual insulin treatment from diagnosis

Genetic
Diagnosis < 6 months OR Specific genetic or clinical diagnosis (MODY, neo-natal diabetes)

Other
Specific primary diagnosis e.g. pancreatitis, pancreatectomy, haemochromatosis, cystic fibrosis, Cushing’s, Acromegaly

Unknown/unclassified
Too early in clinical course to make diagnosis OR uncertainty if specific diagnosis

Non-diabetic hyperglycaemia

IGT ± IFG
Biochemical diagnosis IGT IFG

Gestational Diabetes
Not known diabetes before pregnancy AND Diabetes during pregnancy AND Not diabetes after pregnancy

* In high risk racial groups a cut off of 30 years should be used
Appendix B  Clinical criteria do not precisely classify which insulin treated patients have Type 1 diabetes

Authors: BM Shields, MH Shepherd, N Bhupathi Raju, TJ McDonald, AT Hattersley, Diabetic Medicine, 2010, Volume 27 Issues1;SD4 43.

Background: There are no agreed clinical criteria for classifying Type 1 diabetes (T1D). The Quality Outcomes Framework (QOF) advised using insulin treatment <1y of diagnosis or age at diagnosis <30. T1D is defined by the World Health Organisation as beta-cell destruction leading to absolute insulin deficiency. Therefore, absent endogenous insulin secretion in long duration diabetes robustly defines T1D.

Aim: To determine which clinical criteria can identify T1D, as defined by absolute insulin deficiency, in insulin treated patients.

Methods: We studied 72 insulin-treated adult diabetes patients (>5y duration, insulin within 2y diagnosis). Endogenous insulin secretion was measured using post prandial urinary C-peptide creatinine ratio (UCPCR). Absolute insulin deficiency was defined as <0.2nmol/mmol.

Results: 40/72 patients had T1D as defined by insulin deficiency. QOF criteria were sensitive (93%), but not specific (22%) resulting in an inflated T1D prevalence: 40% QOF “T1D” had measurable C-peptide. Cut-offs derived from Receiver Operating Characteristic curves, had better specificity: age at diagnosis <39 (68% sensitivity, 97% specificity); <=1.5mths to insulin treatment (80% sensitivity, 56% specificity); BMI<29 (78% sensitivity, 56% specificity). Age at diagnosis performed best with cutoff<39y misclassifying fewer patients than QOF criteria (14/72(19%) v 28/72(39%), p=0.01). Combined criteria obtained through regression tree analysis improved the classification modestly (13/72(18%) misclassified).

Conclusion: Compared with the gold standard of chronic insulin deficiency, QOF criteria markedly overestimated the number of insulin treated patients with T1D. Despite being better than QOF, individual or combinations of clinical criteria could not accurately classify all patients. The limitations of clinical criteria have important implications for service planning as well as management of patients.
Appendix C  Acknowledgements

Members of classification of diabetes working group

Professor Kamlesh Khunti (Chair)  Royal College of General Practitioners and the University of Leicester
Professor Simon de Lusignan  University of Surrey
Professor Andrew Hattersley  Peninsula Medical School University of Exeter
Avril Surridge  Patient Representative
Bob Moberley  Patient representative
Liz Allan  NHS Diabetes
Ursula Anderson  NHS Diabetes
Bill O’ Leary  Communications Consultant
Richard Neave  Royal College of General Practitioners
Dr Brian Karet  Diabetes UK and Royal College of General Practitioners
Nathan Moore  Department of Health
Rachel Morton  Department of Health

Funding
The Classification of Diabetes work programme was funded by NHS Diabetes.

Systematic Review
The working group wish to express thanks to Dr Margaret Stone and Janette Camosso-Stefinovic for their invaluable contribution to the Systematic Review.

Database Analysis
The CONDUIT study practices and lead investigators at The Imperial College London, CONDUIT is supported by NIHR and other funding. The QICKD practices and investigators, this study is principally funded by the Health Foundation with additional support from the Edith Murphy Foundation. Funding for this analysis was provided by NHS Diabetes.
Pilot of the Audit Tool
The Working Group wish to express our thanks to the participating practices and also wish to thank the support of the local diabetes team at Royal Surrey Hospital, especially Professor David Russell-Jones, and Henrietta Mulnier, Diabetes Specialist Nurse, for their help and encouragement.

Organisations represented who participated in workshops during the development of this guidance
Association of British Clinical Diabetologists (ABCD)
British Society for Paediatric Endocrinology and Diabetes (BSPED)
Department of Health
Diabetes Inpatient Specialist Nurses (DISN)
Diabetes Research Networks (DRN)
Diabetes UK
Juvenile Diabetes Research Foundation (JDRF)
National Diabetes Nurse Consultant Group
NHS Connecting for Health (CfH)
NHS Diabetes
Primary Care Diabetes Society (PCDS)
Royal College of General Practitioners (RCGP)
Royal College of Nursing (RCN)
Royal College of Physicians (RCP)
User Representatives
Yorkshire and Humber Public Health Observatory (YHPHO)