Received Date : 24-Jan-2011

Revised Date : 27-Apr-2011

Accepted Date : 19-Aug-2011

Article type : Original Article

Corresponding author e-mail id: s.lusignan@surrey.ac.uk

Article: Epidemiology

Miscoding, misclassification and misdiagnosis of diabetes in primary care

S. de Lusignan¹, N. Sadek², H. Mulnier³, A. Tahir¹, D. Russell-Jones^{1,3} and K. Khunti⁴

¹Department of Health Care Policy and Management, Faculty of Management and Law, University of Surrey, Guildford, ²Division of Population Health Sciences and Education, St George's—University of London, London, ³Diabetes and Endocrinology Service, Cedar Centre Royal Surrey County Hospital, Guildford and ⁴Department of Health Sciences, University of Leicester, Leicester, UK

Correspondence to: Simon de Lusignan, Professor of Primary Care and Clinical Informatics, Department of Health Care Policy and Management, Faculty of Management and Law, University of Surrey, Guildford, Surrey GU2 7X, UK. E-mail: s.lusignan@surrey.ac.uk

Running title: Miscoding, misclassification and misdiagnosis of diabetes

This is an Accepted Article that has been peer-reviewed and approved for publication in the *Diabetic Medicine*, but has yet to undergo copy-editing and proof correction. Please cite this article as an "Accepted Article"; doi: 10.1111/j.1464-5491.2011.03419.x

Abstract

Aims To determine the effectiveness of self-audit tools designed to detect miscoding, misclassification and misdiagnosis of diabetes in primary care.

Methods We developed six searches to identify people with diabetes with potential classification errors. The search results were automatically ranked from most to least likely to have an underlying problem. Eight practices with a combined population of 72 000 and diabetes prevalence 2.9% ($n = 2340$) completed audit forms to verify whether additional information within the patients' medical record confirmed or refuted the problems identified.

Results The searches identified 347 records, mean 42 per practice. Pre-audit 20% ($n = 69$) had Type 1 diabetes, 70% ($n = 241$) had Type 2 diabetes, 9% ($n = 30$) had vague codes that were hard to classify, 2% ($n = 6$) were not coded and one person was labelled as having gestational diabetes. Of records, 39.2% ($n = 136$) had important errors: 10% ($n = 35$) had coding errors; 12.1% (42) were misclassified; and 17.0% (59) misdiagnosed as having diabetes. Thirty-two per cent ($n = 22$) of people with Type 2 diabetes ($n = 69$) were misclassified as having Type 1 diabetes; 20% ($n = 48$) of people with Type 2 diabetes ($n = 241$) did not have diabetes; of the 30 patients with vague diagnostic terms, 50% had Type 2 diabetes, 20% had Type 1 diabetes and 20% did not have diabetes. Examples of misdiagnosis were found in all practices, misclassification in seven and miscoding in six.

Conclusions Volunteer practices successfully used these self-audit tools. Approximately 40% of patients identified by computer searches (5.8% of people with diabetes) had errors; misdiagnosis is commonest, misclassification may affect treatment options and miscoding in omission from disease registers and the potential for reduced quality of care.

Keywords automatic data processing, clinical audit, computerized, computing methodologies, diabetes mellitus, diagnostic errors, medical informatics, medical records systems, quality of health care

Abbreviations CONDUIT, Cutting out Needless Deaths Using Information Technology; MIQUEST, Morbidity Information and Export Syntax; QICKD, Quality Improvement in Chronic Kidney Disease trial database

Introduction

There is an international epidemic in diabetes [1] with increased prevalence reported globally [2–4]. However, the reporting of change in prevalence may be undermined if there are errors in routine data which are not addressed. Misclassification matters because it affects treatment choices and risk management. Furthermore, misdiagnosis can have psychological and financial implications, and the combination of miscoding, misclassification and misdiagnosis may undermine measures of the quality of care and research based on routine data [5].

There are known problems with the diagnosis, classification and coding of diabetes [6–8], which have been confirmed in studies using databases drawn from routinely collected data [9]. These databases were from two independent research projects: (1) Cutting out Needless Deaths Using Information Technology (CONDUIT; $n = 221\,958$), a study which explored ethnic disparities in diabetes [10,11] and the longitudinal change in cardiovascular risk management [12]; and (2) Quality Improvement in Chronic Kidney Disease (QICKD; $n = 942\,031$), a newer, much larger study based on a national sample [13,14], which, although primarily a study of quality improvement in chronic kidney disease, includes the investigation of diabetes and other vascular risk factors. These studies concluded that between 10 and 25% of patients with Type 2 diabetes were incorrectly classified as Type 1 diabetes and 5% of those with Type 2 diabetes had no objective evidence of diabetes [9].

For reasons of ethical and practical data management most research databases do not incorporate the whole computerized medical record (CMR). The computerized medical record contains two principal components: (1) coded data—e.g. diagnosis of diabetes, coded as ‘C10’ using the Read classification method in the UK [15]; and (2) free text or narrative record—i.e. the text that the doctor enters into the record and documents, such as hospital and clinic letters. Coded data are relatively easy to process and less likely to compromise a patient’s privacy than free text.

We carried out this study to test the effectiveness of a downloadable toolkit designed to enable practitioners to identify people with diabetes who are miscoded, misclassified or misdiagnosed.

Methods

The Classification of Diabetes (CoD) programme specified the development of audit tools that could be used by general practitioners, to identify patients with incorrectly classified diabetes.

Classification of Diabetes is an initiative of the National Health Service (NHS) Diabetes and the Royal College of General Practitioners (RCGP) in England and information can be accessed online [16]. The Classification of Diabetes programme includes the development of audit tools that practitioners or localities can use to identify patients with classification problems. We followed the model we have previously used to identify patients suitable for referral to a new psychological therapy service, which consisted of data extraction queries and a step-by-step user guide for most of the brands of primary

care computerized medical record systems [17]. A flow chart provides an overview of the process (Fig. 1).

We incorporated the naming conventions used in the Classification of Diabetes in this paper. These include definitions of miscoding, misclassification and misdiagnosis, as well as a practical algorithm for classifying diabetes. (1) Miscoding occurs when the wrong code is applied, generally the use of a non-specific code. For example, when use of a non-specific diabetes code is used, it does not allow the type of diabetes to be precisely determined. (2) Misclassification is defined as when a case is incorrectly classified into a category to which the subject does not belong. (3) Misdiagnosis is the allocation of an incorrect diagnosis; diagnosing diabetes in someone who does not have the condition. The algorithm provides practical guidance on how to classify diabetes according to World Health Organization (WHO) categories. However, it adds an 'Unclassified' group for patients who are early in the diagnostic process or in whom there is genuine uncertainty; and groups with impaired glucose tolerance or with gestational diabetes into 'Non-diabetic hyperglycaemia'. It also uses the classification 'other' to include the secondary causes of diabetes.

We carried out this audit in eight practices, five in Surrey and three in south-west London. The practices had a combined list size of 72 000 [18] and a mean of 9000 patients (median 10 043) per practice. However, the Surrey practices were larger (mean 12 931) compared with the London practices (mean 2448). The practices had all created a disease register of people with diabetes, as part of pay-for-performance (P4P) quality targets [19]. These disease registers contained a total of 2340 people with diabetes, representing an overall prevalence of 3.2% (range 2.9–3.9%). The auditors were trained by one of the authors (SdeL), including jointly running through cases until the auditor was happy to continue the process; SdeL met and reviewed with the auditor any problematic cases and discussed these with a co-author (KK). The audits were carried out in two practices by general practitioners who ran their practice diabetes clinic, in one by a specialist diabetes nurse, and in two by practice nurses who were trained and ran their practice-based clinics. In three practices, the audits were initiated by general practitioner diabetes leads, but the audit was carried out by a final year medical student (NS).

We developed six simple Morbidity Information Query and Export Syntax (MIQUEST) computer searches to explore whether it was feasible for these to be run in practice. MIQUEST is a Department of Health-sponsored data extraction tool, capable of extracting data from different brands of general practitioner computer system. We ran these queries in 'local' mode within individual practices, so that it produced simple data tables of patients requiring notes reviewed for the practice to process. The MIQUEST query processor and general practitioner computerized record systems cannot carry out the sophisticated data processing carried out in the original study [5], so we developed searches,

technically called queries, using simple logic to identify the groups of patients we were interested in (Fig. 2). We accepted that there may be some overlap between the queries.

The six queries extracted diagnostic data (disease code and date) and drug therapy codes (relevant drug code and date) to look for inconsistencies between diagnosis and therapy or biochemistry. The first three queries identify people with misclassification of Type 1 diabetes. They look for people with the Type 1 diabetes disease code not prescribed insulin; where insulin and an oral anti-diabetes drug (OAD) are prescribed; and where oral anti-diabetes drugs are started before insulin and insulin is not started within 6 months of diagnosis. The fourth query explores misclassification of Type 2 diabetes by extracting the Type 2 diabetes disease code and then looking for prescription codes for insulin within 6 months of diagnosis. The fifth query looks to identify misdiagnosis of Type 2 diabetes by extracting the disease code for Type 2 diabetes, oral anti-diabetes drug and Insulin data, plasma glucose and HbA_{1c} data. These data are sorted in two ways: firstly, to identify people with a disease code compatible with Type 2 diabetes, who are not taking an oral anti-diabetes drug or insulin, and who have not ever had an HbA_{1c} ≥ 48 mmol/mol (6.5%) or plasma glucose ≥ 7.0 mmol/l; secondly, the query identifies the people taking an oral anti-diabetes drug who lack a diagnosis of diabetes. The final query flags people with vague diabetes codes. The precise code ranges used are listed within the queries, which can be downloaded from our website (<http://www.clininf.eu/cod>).

When we tested the MIQUEST searches, we found that the interface between MIQUEST and the computer system is implemented in subtly different but important ways in the various brands of computerized medical record systems. For example, the MIQUEST manual [20] says that you can request maximum values, for example, highest plasma glucose. However, this feature is not implemented in one of the general practitioner computerized medical record systems we worked with. We therefore had to produce queries that listed all patients with a condition, treatment or pathology result and then reorder them within a spread sheet (Microsoft Excel) to identify the high-risk patients.

The next steps of the process involved developing a user-friendly method of data sorting and an audit form to collect data about the cases identified. We developed a Microsoft Excel macro which allowed the output files generated by the queries to be placed in a single folder and be sorted from most likely to least likely to have coding or classification problems. We did not precisely specify the number of cases each practice should review. We created an illustrated step-by-step user guide and an audit form and placed these online [21]. The audit tool was developed as part of the Classification of Diabetes programme and was initially piloted in two practices. An audit form was completed for each case reviewed. The audit was designed to collect the minimum amount of data required to clarify an individual patient's diagnosis. The audit collected the most recent data available for a patient as well as any data collected at the point of diagnosis or at the earliest record of having diabetes.

Statistical tests

We used Pearson χ^2 -test to compare whether there were any statistical differences in proportion between groups. We used an independent samples *t*-test to see if there are any differences in age, plasma glucose, HbA_{1c} or BMI between those found to be misdiagnosed and those with diabetes and a paired sample *t*-test to compare any change in continuous variables between diagnosis and latest measurement. We conducted our analysis using SPSS (PASW statistics) version 18 (SPSS Inc., Chicago, IL, USA).

Ethical considerations

The findings of the study of coded data from the CONDUIT study and QICKD trial data led NHS Diabetes to roll out a national programme to improve the quality of the classification of diabetes [22] in line with WHO recommendations [23]. Identifiable data were only held at practice level, with anonymized data held at St George's and only capable of being re-identified within contributing practices. This audit was a test of implementing these national recommendations by volunteer practices with direct responsibility of the care of their patients, and consistent with the General Medical Council guidance to participate in local audit [24] and National Research Ethics Service definitions [25] of clinical audit.

Results

Cases audited

Audit forms were completed for 347 patients representing 14.8% (347/2340) of the people labelled as having diabetes in the practice computerized medical record system. There were 162 women with a median age of 65.5 years (range 15–97 years; interquartile range 49–76 years). The corresponding data for men ($n = 185$) were: median age 63 years (range 16–95 years, interquartile range 54–71 years). The audit form took between 10 and 40 min to complete. A small number of cases were complex and required extensive review of the whole record and discussion with authors SdeL and KK.

All types of diabetes were identified by the audit, including cases not included in the pay-for-performance disease register. Most of the cases identified were Type 2 diabetes (69.5%, $n = 241$) and Type 1 diabetes (19.9%, $n = 69$). Eleven per cent ($n = 37$) of patients identified by the searches would not have been included in the pay-for-performance disease registers. This was either because they use vague codes not readily mapped to a type of diabetes or because they were identified from being prescribed treatment for diabetes (Table 1).

Miscoding, misclassification and misdiagnosis of diabetes

As a result of the case review, at least 16% ($n = 59$) of the patients diagnosed with diabetes were found to be incorrectly diagnosed and did not have the condition. These 59 comprised 57 without diabetes (13 of whom had impaired glucose tolerance) and two with gestational diabetes (Table 2). There were two cases within the audit where the patients concerned had no objective record of diabetes within their record other than a single blood glucose reading; these were 16.4 mmol/l and 20.9 mmol/l. These results superficially meet the diagnostic criteria for diabetes [22].

We identified misclassification with people with Type 2 diabetes being classified as having Type 1 diabetes; misdiagnosis of Type 2 diabetes when patients actually did not have diabetes; and patients who had vague diabetes codes or prescribed medications for diabetes, most of whom had Type 2 diabetes. Thirty-two per cent of patients (22 out of 69) flagged with Type 1 diabetes actually had Type 2 diabetes (Table 3). Eleven patients coded as having Type 2 diabetes had Type 1 diabetes and 48 patients coded with Type 2 diabetes did not have diabetes at all. Of the patients with a vague diagnosis code for their diabetes (none of whom could not be readily classified as Type 1 diabetes or Type 2 diabetes), four did not have diabetes, six had Type 1 diabetes and 15 had Type 2 diabetes. Five cases of secondary diabetes were identified in patients with a pre-audit diagnosis of Type 2 diabetes. In the four cases where a pre-audit diagnosis of Type 1 diabetes turned out not to be diabetes: two cases were patients who had pancreatic transplants, although these may be patients who continue to be reviewed; one a very obese person who had been on insulin at one stage but there was no evidence of diabetes (there were no elevated plasma glucose or glycated haemoglobin results in his records, although a strong family history was noted and we cannot rule out that this person did not have Type 2 diabetes now controlled by diet.) The final case appeared to be a coding error.

Overall impact of the audit

Overall, the effect of the audit was to change the diagnosis of approximately 40% ($n = 136$) of the patients identified by the tool. This comprised 5.8% (136/2340) of the patients on the diabetes registers across these practices. There was no change in 60.8% of the patients ($n = 211$); 17% (59) had a misdiagnosis, based on non-diabetic levels of plasma glucose and HbA_{1c}, and no need for ongoing therapy; 12.2% (42) were misclassified and 10.1% (35) had been miscoded.

Which searches were most useful

All six of the searches in the audit identified people who required change (Table 4). The searches which identified the largest proportion of people requiring change in coding were those labelled as having Type 2 diabetes. People with Type 1 diabetes requiring a change in coding were more often identified through the process of completing the audit form.

Pointers towards patients more likely to require change

We explored whether problems in coding, classification or diagnosis in patients' records were different in terms of age, BMI or glycaemic control. The mean ages of patients who were miscoded, misclassified or misdiagnosed and those who were not were similar (61.8 vs. 61.3 years). People with miscoding, misclassification and misdiagnosis had similar BMI (28.5 vs. 28.1 kg/m²) at the time of diagnosis. However, whilst there was no difference in HbA_{1c} at the time of diagnosis [HbA_{1c} = 62 mmol/mol (7.8%) vs. 61 mmol/mol (7.7%) for no change vs. change groups], in the latest results HbA_{1c} was significantly lower in the group whose coding was changed [HbA_{1c} = 56 mmol/mol (7.3%), SD = 2.14 vs. HbA_{1c} = 50 mmol/mol (6.7%), SD = 2.14; *t*-test, *P* = 0.019].

Comparing patients misdiagnosed with a combined misclassified and miscoded group

We also compared those misdiagnosed (i.e. patients we thought did not have diabetes) with the misclassified and miscoded as a single group and discovered that the misdiagnosed group have a lower HbA_{1c} and plasma glucose, although there was no difference in their age or BMI at diagnosis or at their latest reading. The patients who were misclassified or miscoded had a mean HbA_{1c} of 78 mmol/mol (9.3%) (SD 8.5), whereas those misdiagnosed had an HbA_{1c} of 40 mmol/mol (5.8%) (SD 1.7) at diagnosis. The most recent HbA_{1c} values were also significantly different [62 mmol/mol (7.8%) or vs. 34 mmol/mol (5.3%), respectively]. Plasma glucose at diagnosis was 12.8 mmol/l (SD 6.0) vs. 8.5 mmol/l (SD 3.2). These differences are all highly statistically significant (*t*-test, *P* = 0.003 for earliest HbA_{1c} and *P* < 0.001 for plasma glucose and latest HbA_{1c}).

Over the time between first recording and the most recent reading, the misdiagnosed groups maintained much lower HbA_{1c} readings [earliest HbA_{1c} 40 mmol/mol (5.8%) (SD 1.7), latest 34 mmol/mol (5.3%) (SD 1.0), *P* = 0.001].

Variation in cases found by practice

Cases were identified in all practices; the median number of cases audited per practice was 44 (interquartile range 36.5–55.3) and the median number where a change was needed was 17.0 (interquartile range 10.8–18.8). All practices had at least one misdiagnosed case, all but one identified a case of misclassification and all but two found examples of miscoding (Table 4).

Discussion

Principal findings

This audit shows that there is scope to improve diabetes data quality and that computer searches may have a role in achieving such a rise in data quality. The audit also demonstrates how self-directed audits can be downloaded and successfully conducted in practice and validates the findings of previous studies using research databases. Approximately 40% of the cases identified through the

audit had a coding change as a result, with misdiagnosis being the largest group. Two patients who had 'diabetic' blood test results subsequently have no features of diabetes.

Implications of the findings

Some practices are able to run this self-directed audit toolkit, and this approach to clinical audit has a place among the armamentarium of quality improvement tools. There is scope to improve the quality of care of people currently labelled as having diabetes; some may not have diabetes; others are misclassified and may not be managed according to correct guidance; and people who are miscoded will not be included in the pay-for-performance register so will not be automatically flagged or recalled by most general practitioner computerized medical record systems. Where they are not automatically recalled (such as secondary causes of diabetes), practice recall systems need to be put in place to ensure they receive quality care. Where patients appear to be very well controlled and have consistently low HbA_{1c}, practices should consider the possibility that the patient does not have diabetes. There should be a high level of suspicion where a patient was diagnosed on a single blood test.

Comparison with the literature

This audit was conducted because a previous systematic review identified few studies of classification problems in diabetes [9]. The prevalence of diabetes based on the participating practices' pay-for-performance disease registers is 3.2% and within 0.1% of a national survey of pay-for-performance prevalence [26]. The pay-for-performance register is likely to underestimate prevalence because some patients with diabetes were coded using vague codes that are not included in the disease register. This partially explains why studies of routine data that are inclusive of all relevant codes report a higher prevalence, approximately 4.0% [27–29].

Computer prompts could potentially improve quality and coding of diabetes. Randomized controlled trials in family practice have found reminders to test for diabetes have improved detection [30] and prompts to improve primary prevention are effective [31]. Patients who are included in the pay-for-performance disease register will be subject to reminders and prompts, albeit in varying formats on the different brands of computerized medical record system [32], whereas those with diabetes not included in this register will not be. It is plausible that exclusion from the pay-for-performance disease register contributes to lower standards of glycaemic control.

HbA_{1c} has been proposed as a better diagnostic tool than fasting glucose or oral glucose tolerance tests [33]. Where a diagnosis of diabetes has been made on the basis of a single plasma glucose test, and HbA_{1c} is below the levels proposed for the diagnosis of diabetes, we suggest that practitioners consider challenging the diagnosis.

Limitations of the study

The limitations of working with routine data, which is inevitably incomplete, are well known [34]. The clinicians conducting the audit often opportunistically noted problems with the coding and classification of diabetes, which were not necessarily those intended to be highlighted by the search. We did not manually search all the records so are unable to comment on the sensitivity of this method. Whilst we found a MIQUEST-naïve medical student could complete the searches, it is possible that some practices may lack someone with the skills to run these searches. These practices may be able to get support from a local data quality or audit facilitator.

Call for further research

Further study would enable us to develop and improve these audit tools further and explore false positives and cases missed. Process evaluation would provide insight into the usability of this toolkit [35].

Conclusions

Problems with the coding classification and diagnosis of diabetes seen in a study of coded data from research databases are present in routine data and practices should consider running the audit (available at <http://www.clininf.eu/cod>). However, over half the patients identified required no change and individual patient problems were often complex and sometimes not the ones the search was intended to find. Every person with diabetes who attends for review should be critically reviewed to check whether they are correctly classified and coded.

Competing interests

SdeL is informatics work-task lead for the NHS Diabetes Classification of Diabetes task group; KK is lead for the same group; NS, HM, AT and DRJ have nothing declare.

Acknowledgements

We acknowledge participating practices, NHS Diabetes and RCGP Classification of Diabetes project team. We also acknowledge Mark Bradley of Mehdi-Ward for developing the query-sorting macro, Andre Ring for collecting the data and Elena Crecan and Anthonis Ntasioudis for assembling the audit outputs.

References

1. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. *Nature* 2001; **414**: 782–787.
2. Pan C, Shang S, Kirch W, Thoenes M. Burden of diabetes in the adult Chinese population: a systematic literature review and future projections. *Int J Gen Med* 2010; **3**: 173–179.
3. Abubakari AR, Lauder W, Jones MC, Kirk A, Agyemang C, Bhopal RS. Prevalence and time trends in diabetes and physical inactivity among adult West African populations: the epidemic has arrived. *Public Health* 2009; **123**: 602–614.
4. Lusignan S, Sismanidis C, Carey IM, DeWilde S, Richards N, Cook DG. Trends in the prevalence and management of diagnosed type 2 diabetes 1994–2001 in England and Wales. *BMC Fam Pract* 2005; **6**: 13.
5. Stone MA, Camosso-Stefinovic J, Wilkinson J, de Lusignan S, Hattersley AT, Khunti K. Incorrect and incomplete coding and classification of diabetes: a systematic review. *Diabet Med* 2010; **27**: 491–497.
6. Rollason W, Khunti K, de Lusignan S. Variations in the recording of diabetes diagnostic data in primary care computer systems: implications for the quality of care. *Inform Prim Care* 2009; **17**: 113–119.
7. Bagheri A, Sadek A, Chan T, de Lusignan S. Using surrogate markers in primary electronic patient record systems to confirm or refute the diagnosis of diabetes. *Inform Prim Care* 2009; **17**: 121–129.
8. 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**: 63–64.
9. de Lusignan S, Khunti K, Belsey J, Hattersley A, van Vlymen J, Gallagher H *et al*. A method of identifying and correcting miscoding, misclassification and misdiagnosis in diabetes: a pilot and validation study of routinely collected data. *Diabet Med* 2010; **27**: 203–209.
10. Gray J, Millett C, Saxena S, Netuveli G, Khunti K, Majeed A. Ethnicity and quality of diabetes care in a health system with universal coverage: population-based cross-sectional survey in primary care. *J Gen Intern Med* 2007; **22**: 1317–1320.
11. Millett C, Gray J, Saxena S, Netuveli G, Khunti K, Majeed A. Ethnic disparities in diabetes management and pay-for-performance in the UK: the Wandsworth Prospective Diabetes Study. *PLoS Med* 2007; **4**: e191.
12. Murray J, Saxena S, Millett C, Curcin V, de Lusignan S, Majeed A. Reductions in risk factors for secondary prevention of coronary heart disease by ethnic group in south-west London: 10-year longitudinal study (1998–2007). *Fam Pract* 2010; **27**: 430–438.
13. de Lusignan S, Gallagher H, Chan T, Thomas N, van Vlymen J, Nation M *et al*. The QICKD study protocol: a cluster randomised trial to compare quality improvement interventions to lower systolic BP in chronic kidney disease (CKD) in primary care. *Implement Sci* 2009; **4**: 39.

14. de Lusignan S, Tomson C, Harris K, van Vlymen J, Gallagher H. Creatinine fluctuation has a greater effect than the formula to estimate glomerular filtration rate on the prevalence of chronic kidney disease. *Nephron Clin Pract* 2010; **117**: c213–c224.
15. de Lusignan S. Codes, classifications, terminologies, and nomenclatures: definition, development and application in practice: a theme of the European Federation for Medical Informatics Primary Care Informatics Working Group (EFMI PCI WG). *Inform Prim Care* 2005; **13**: 65–69.
16. NHS Diabetes/RCGP. *Classification of Diabetes (CoD) Programme*. Available at http://www.diabetes.nhs.uk/our_work_areas/classification_of_diabetes Last accessed XX Xxxxx XXXX.
17. Clinical Informatics, St George's–University of London. *MIQUEST User Manuals and Queries for Data Extraction: IAPT Study*. Available at <http://www.clininf.eu/iapt.html> Last accessed XX Xxxx XXXX.
18. Prescription Pricing Authority (PPA). *Data Released under the Freedom of Information Act, 2010*. Available at [http://www.ppa.org.uk/foidocs/responses/FOI_Request\(344010\).pdf](http://www.ppa.org.uk/foidocs/responses/FOI_Request(344010).pdf) Last accessed XX Xxxxx XXXX.
19. NHS Information Centre for Health and Social Care. *Statistics and Data Collections: Practice Level Data Tables*. Available at <http://www.ic.nhs.uk/statistics-and-data-collections/supporting-information/audits-and-performance/the-quality-and-outcomes-framework/qof-2009-10/data-tables/practice-level-data-tables> Last accessed XX Xxxx XXXX.
20. NHS Connecting for Health. *MIQUEST Specification Download Document Version 4.2.3*. Available at <http://www.connectingforhealth.nhs.uk/systemsandservices/data/miquest/docs> Last accessed XX Xxxxx XXXX.
21. Clinical Informatics. *MIQUEST User Manuals and Queries for Data Extraction: Diabetes*. Available at <http://www.clininf.eu/diabetes.html> Last accessed XX Xxxxx XXXX.
22. NHS Diabetes. *Getting it Right: Improving the Classification, Diagnosis and Coding of Diabetes*. Fact sheet 34. Leicester: NHS Diabetes, 2011. Available at http://www.diabetes.nhs.uk/publications_and_resources Last accessed XX Xxxxx XXXX.
23. Alberti K, Zimmet P. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998; **15**: 539–553.
24. General Medical Council (GMC). *Guidance on Good Practice: Confidentiality Guidance: Circumstances in which Patients give Implied Consent to Disclosure*. Available at http://www.gmc-uk.org/guidance/ethical_guidance/confidentiality_24_35_disclosing_information_with_consent.asp Last accessed XX Xxxx XXXX.
25. NRES. Defining Research Leaflet. NHS National Patient Safety Agency (NPSA). Ref. 0987; London: National Research Ethics Service, 2008. Available at <http://www.nres.npsa.nhs.uk/applications/guidance/research-guidance> Last accessed XX Xxxxx XXXX.

26. Millett C, Car J, Eldred D, Khunti K, Mainous AG 3rd, Majeed A. Diabetes prevalence, process of care and outcomes in relation to practice size, caseload and deprivation: national cross-sectional study in primary care. *J R Soc Med* 2007; **100**: 275–283.
27. Gray J, Orr D, Majeed A. Use of Read codes in diabetes management in a south London primary care group: implications for establishing disease registers. *Br Med J* 2003; **326**: 1130.
28. New JP, Middleton RJ, Klebe B, Farmer CK, de Lusignan S, Stevens PE *et al*. Assessing the prevalence, monitoring and management of chronic kidney disease in patients with diabetes compared with those without diabetes in general practice. *Diabet Med* 2007; **24**: 364–369.
29. González EL, Johansson S, Wallander MA, Rodríguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996–2005. *J Epidemiol Community Health* 2009; **63**: 332–336.
30. Kenealy T, Arroll B, Petrie KJ. Patients and computers as reminders to screen for diabetes in family practice. Randomized controlled trial. *J Gen Intern Med* 2005; **20**: 916–921.
31. Holt TA, Thorogood M, Griffiths F, Munday S, Friede T, Stables D. Automated electronic reminders to facilitate primary cardiovascular disease prevention: randomised controlled trial. *Br J Gen Pract* 2010; **60**: e137–e143.
32. Debar S, Kumarapeli P, Kaski JC, de Lusignan S. Addressing modifiable risk factors for coronary heart disease in primary care: an evidence-base lost in translation. *Fam Pract* 2010; **27**: 370–378.
33. International Expert Committee. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. *Diabetes Care* 2009; **32**: 1327–1334.
34. 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**: 253–263.
35. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. *Qual Saf Health Care* 2003; **12**: 40–46.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Audit form.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than for missing material) should be directed to the corresponding author for the article.

Figure 1 Audit overview flow chart. MIQUEST, Morbidity Information Query and Export Syntax.

Figure 2 Six queries to identify patients with miscoded, misclassified and misdiagnosed diabetes.

Three queries to look at misclassification of Type 1 diabetes

1. Patients with a diagnosis of type 1 diabetes not prescribed insulin
2. Patients with a diagnosis of type 1 diabetes on insulin currently taking an oral anti-diabetic medicine other than metformin. Patients with type 1 diabetes should not routinely be concurrently prescribed an oral anti-diabetic medicine, an exception being metformin for weight loss
3. Patients with a diagnosis of type 1 diabetes who started their oral anti-diabetic medicine before insulin will have started insulin within 6 months of diagnosis .

One query to look at misclassification of Type 2 diabetes

4. Patients who have required insulin from within 6 months of diagnosis. Patients prescribed insulin but who do not have a diagnosis of diabetes are also detected by this query

One query to look at misdiagnosis of Type 2 diabetes

5. Patients with type 2 diabetes not prescribed insulin or an oral anti-diabetic medicine (excluding metformin) or had no abnormal tests—plasma glucose and HbA_{1c}—are unlikely to have diabetes. Patients prescribed an oral anti-diabetic medicine (excluding metformin) but who do not have a diagnosis of diabetes are also detected by this query

Miscoding in diabetes

6. Some patients with diabetes are miscoded. Their disease is only coded with vague codes (e.g. C100z 'Diabetes without mention of complications'), which inform the patient has diabetes but cannot be classified

Table 1 Cases identified by running the searches listed in Fig. 2

	Audit cases	
	<i>n</i>	%
Type 1 diabetes	69	19.9
Type 2 diabetes	241	69.5
Vague/unclassified	30	8.6
Not coded	6	1.7
Gestational	1	0.3
Total	347	100

Pre-audit classification		Post-audit classification															
	Pre-audit Total	Type 1 diabetes		Type 2 diabetes		Secondary		Unclassified		Patients with diabetes Subtotal		Non diabetic hyperglycaemia		Not diabetes	Patients without diabetes Subtotal		
										Gestation al	Impaired glucose tolerance						
	<i>n</i>	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%		
Type 1 diabetes	69	35	50.7	22	31.9	2	2.9	5	7.2	64	92.8	0	0.0	0	0.0	5	7.2
Type 2 diabetes	241	11	4.6	174	72.2	5	2.1	3	1.2	193	81.0	0	0.0	13	5.4	35	14.5
Vague/unclassified	30	6	20.0	15	50.0	0	0.0	3	10.0	24	83.3	2	6.7	0	0.0	4	13.3
Not coded	6	1	16.7	3	50.0	0	0.0	2	33.3	6	100	0	0.0	0	0.0	0	0.0
Gestational	1	0	0.0	1	100	0	0.0	0	0.0	1	100	0	0.0	0	0.0	0	0.0
Total	347	53	15.3	215	62.0	7	2.0	13	3.7	291	83.9	2	0.6	13	3.7	41	11.8

Table 2 Cross-tabulation of pre- and post-audit types of diabetes

Table 3 The cases and problems identified by each of the six searches

Query No.	Problem type the query intended to identify	No change		Miscoding		Misclassification		Misdiagnosis		Total	
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
1	Misclassification type 1 diabetes	15	7.1	1	2.9	8	19.0	3	5.1	27	7.8
2	Misclassification type 1 diabetes	24	11.4	12	34.3	17	40.5	1	1.7	54	15.6
3	Misclassification type 1 diabetes	4	1.9	1	2.9	2	4.8	0	0.0	7	2.0
4	Misclassification type 2 diabetes	56	26.5	7	20.0	11	26.2	5	8.5	79	22.8
5	Misdiagnosis type 2 diabetes	43	20.4	0	0.0	2	4.8	40	67.8	85	24.5
6	Miscoding in diabetes	69	32.7	14	40.0	2	4.8	10	16.9	95	27.4
Total		211	100	35	100	42	100	59	100	347	100

Table 4 Changes as a result of the audit by practice

Practice	Location	List (rounded)	Audit cases identified Cases as % of list size		Diabetes pay for performance register Diabetes prevalence		Change following audit % pay for performance diabetes register		Miscoded % pay for performance diabetes register		Misclassified % pay for performance diabetes register		Misdiagnosed % pay for performance diabetes register	
		N	n	%	n	%	n	%	n	%	n	%	n	%
1	Surrey	11 000	53	0.5	438	3.9	17	3.9	10	2.3	5	1.1	2	0.5
2	Surrey	12 000	41	0.3	373	3.2	10	2.7	0	0.0	8	2.1	2	0.5
3	Surrey	9000	47	0.5	254	2.9	21	8.3	6	2.4	0	0.0	15	5.9
4	Surrey	15 000	64	0.4	477	3.2	33	6.9	3	0.6	11	2.3	19	4.0
5	Surrey	18 000	26	0.1	525	2.9	18	3.4	0	0.0	9	1.7	9	1.7
6	London	2500	40	1.7	91	3.9	11	12.1	3	3.3	3	3.3	5	5.5
7	London	2500	14	0.5	101	3.9	9	8.9	3	3.0	5	5.0	1	1.0
8	London	2500	62	2.6	81	3.3	17	21.0	10	12.3	1	1.2	6	7.4
	TOTAL	72 000	347	0.5	2340	3.2	136	5.8	35	1.5	42	1.8	59	2.5