Research: Care Delivery

Glucose test provenance recording in UK primary care: was that fasted or random?

A. P. McGovern[1], H. Fieldhouse[1], Z. Tippu[1], S. Jones[1], N. Munro[1] and S. de Lusignan[1,2]

[1]Department of Clinical and Experimental Medicine, University of Surrey, Guildford and [2]Clinical Innovation and Research Centre, Royal College of General Practitioners, London, UK

Abstract

Aims   To describe the proportion of glucose tests with unrecorded provenance in routine primary care data and identify the impact on clinical practice.

Methods   A cross-sectional analysis was conducted of blood glucose measurements from the Royal College of General Practitioner Research and Surveillance Centre database, which includes primary care records from >100 practices across England and Wales. All blood glucose results recorded during 2013 were identified. Tests were grouped by provenance (fasting, oral glucose tolerance test, random, none specified and other). A clinical audit in a single primary care practice was also performed to identify the impact of failing to record glucose provenance on diabetes diagnosis.

Results   A total of 2 137 098 people were included in the cross-sectional analysis. Of 203 350 recorded glucose measurements the majority (117 893; 58%) did not have any provenance information. The most commonly reported provenance was fasting glucose (75 044; 37%). The distribution of glucose values where provenance was not recorded was most similar to that of fasting samples. The glucose measurements of 256 people with diabetes in the audit practice (size 11 514 people) were analysed. The initial glucose measurement had no provenance information in 164 cases (64.1%). A clinician questioned the provenance of a result in 41 cases (16.0%); of these, 14 (34.1%) required repeating. Lack of provenance led to delays in the diagnosis of diabetes [median (range) 30 (3–614) days].

Conclusions   The recording of glucose provenance in UK primary care could be improved. Failure to record provenance causes unnecessary repeated testing, delayed diagnosis and wasted clinician time.

Diabet. Med. 00, 000–000 (2016)

Introduction

Blood glucose fluctuates considerably, with peaks after meals or intake of sugar and carbohydrates. In diseases of blood glucose control, particularly diabetes mellitus, the measurement of blood glucose is fundamental to making the diagnosis, but knowing the provenance of the glucose sample is essential to be able to interpret the result. Diabetes diagnosis can alternatively be made using glycated haemoglobin (HbA_{1c}) measurement [1]. In practice HbA_{1c} is now used for diagnosis in most people.

Primary care internationally uses computerized medical record systems in which key data are coded [2]. The UK currently uses the Read code system to record data, including laboratory test results. Inputting of these data from the laboratory is an automated process. Laboratory requests are almost exclusively made online. The results are sent back automatically and then the clinician files the results from a suspense file into the clinical record. The laboratory assigns a Read code to the result that cannot be readily changed. For blood glucose the particular Read code used defines the provenance of the sample to be recorded.

Concerns have previously been raised regarding the failure to use the correct Read code to document whether recorded glucose values are fasted samples [3]. We have also previously reported that the majority of recorded glucose tests used to follow up women with gestational diabetes in primary care did not specify the type of glucose test used; 95 samples out of 146 (65%) [4].

The provenance of data is an aspect of data quality. It is defined as 'how data came to be' [5]. A blood glucose record may have come about from a number of sources, laboratory entered or directly coded into the clinical record by the

clinician. We searched a large database for all glucose Read codes of different provenances; this search incorporated codes used for laboratory results, clinician-entered glucometer values and glucose tolerance tests performed at centres other than the registered primary care centre.

Based on these previous observations, we hypothesized that the majority of glucose samples recorded in primary care in the UK would not be coded according to the type of test performed. We used data collected for the Royal College of General Practitioner Research and Surveillance Centre (RCGP RSC) database to quantify the coding of blood glucose test provenance and to make recommendations for future practice based on these data. We also performed a clinical audit in a primary care practice to identify the impact of failing to record glucose provenance. To the best of our knowledge, this has not been investigated previously.

Methods

We report a mixed methods study with two components: firstly, we report a cross-sectional analysis of serum glucose measurements in a large primary care population to identify the proportion of tests with no provenance information and, secondly, we report a clinical audit of the impact of deficiencies in glucose test provenance recording in all patients with diabetes in an individual UK primary care practice.

Cross-sectional analysis

A cross-sectional analysis, using data from a large primary care-based population, was conducted to report the clinical codes used to record the type of glucose sample in the electronic patient record.

Anonymized data, collected routinely by the RCGP RSC, were used to define the cohort. These data comprise the electronic patient records from patients registered at >100 primary care practices across England and Wales. These practices have been selected by the RCGP RSC to provide a representative sample of urban, suburban and rural practices. The primary aim of this data collection is for disease surveillance of the incidences of influenza-like illness and other infectious diseases [6]; however, the dataset has been

used for a number of epidemiological studies looking at both infectious diseases and other conditions [6–9]. Data are automatically uploaded from the patient record to the database from all the included practices via regular data uploads.

We identified all the blood glucose tests recorded during 2013, this included both laboratory results and tests performed using glucometers. We included all patients in the RCGP RSC database in this search. We did not explore pregnancy status. Blood glucose tests were grouped by provenance; fasting glucose, diagnostic oral glucose tolerance test (OGTT) glucose (tests coded as occurring at 2 h as part of an OGTT), random glucose, other provenance, and no provenance specified. The complete codes list included for each provenance is shown in Appendix S1.

The distribution of glucose values from tests with no provenance was compared with the distribution of fasting glucose, random glucose and glucose tolerance test values using a two-sample Kolmogorov–Smirnov test (comparison of the cumulative distributions). The kurtosis and skew for each distribution are also reported for comparison. For this analysis, glucose values were cleaned, removing non-numeric and missing data. All glucose values > 36 were assumed to be recorded in mg/dl rather than mmol/l. These values were converted into mmol/l by dividing by 18.02. Glucose values < 1 mmol/l and > 36 mmol/l were excluded. Data outside these ranges had a high proportion of entries likely to be the result of data inputting errors and were therefore considered unreliable. The number of missing results and erroneous entries are reported.

We also report the proportion of test results which met the diagnostic criteria for diabetes and the diagnostic criteria for impaired fasting glucose (6.1–6.9 mmol/l) and impaired glucose tolerance (7.8–11.0 mmol/l) [10]. All statistical analyses were performed using the software package R version 2.15.2.

Clinical audit

A clinical audit of all the patients with a diagnosis of diabetes in a single primary care practice in England was undertaken to identify whether missing glucose provenance information had any impact on the initial diagnosis of diabetes.

People with diabetes were identified using the practice's disease register of people with diabetes, created as part of pay-for-performance quality targets. All patients with diabetes at the practice are included on this register and the practice has previously undertaken a clinical audit of its diabetes register to identify miscoded, misclassified and misdiagnosed patients, with ~6% of people with diabetes requiring correction of coding errors [11]. After these corrections were made, this practice provided an accurate sample of people with diabetes.

The clinical records of all the registered people with diabetes were reviewed by a clinician. The date of diagnosis

was identified and the related investigations were recorded along with the recorded glucose provenance information (recorded as free text and clinical Read code). Any clinician comments regarding the provenance of glucose test and subsequent clinical actions were also recorded using a prepared audit *pro forma*.

We report the proportion of people with diabetes for whom the diagnosis was made using serum glucose measurements, the proportion of first and second glucose tests where glucose provenance was not recorded or coded, and the negative clinical impacts of failure to record glucose provenance. The negative impacts investigated were: number of times clinicians contacted patients to attempt to identify test provenance (as documented in the clinical record); the number of repeat tests undertaken because the test provenance could not be determined; and delays in diagnosis. The delay in diagnosis was defined as the time interval between a first test suggestive of diabetes ($\geq$ 7.0 mmol/l), where no provenance information was recorded, and the time of confirmation of diabetes by a test with recorded glucose provenance or confirmation of the diagnosis via an alternative method.

Ethical considerations

The RCGP RSC approved the conduct of the present study using RCGP data. The audit component of the study was to identify the quality of implementation of accepted recommendations for the diagnosis of diabetes [1,10] by a volunteer practice with direct responsibility for the care of their patients, with collected data used to improve local patient care. The audit was consistent with the General Medical Council guidance for participation in audit [12] and National Research Ethics Service definition [13] of clinical audit.

Results

Cross-sectional analysis

The primary care records of 2 137 098 people were included for analysis. From these records we identified 203 350 blood glucose measurements taken between 1 January 2013 and 31 December 2013 inclusive. A total of 146 901 (6.87%) people had one or more glucose tests coded.

No provenance was specified for the majority of glucose tests (Table 1). The majority of codes were recorded with a valid numerical value for fasting, diagnostic oral glucose tolerance test (OGTT), random, and no provenance glucose values. The most frequently used codes to record glucose measurement were 44g., 'plasma glucose level' (76 905 tests; 34.5% of all recorded tests) and 44g1., 'fasting plasma glucose level' (53 252 tests; 23.9% of all tests).

The distribution of glucose values for tests with unknown provenance most closely resembles that of fasted samples, suggesting the majority of these samples were taken with the patient having fasted (Fig. 1 and Table 2); however, the two distributions are significantly different; two-sample Kolmogorov–Smirnov statistic for comparison of the cumulative distribution of fasting glucose tests with those of unknown provenance D=0.3744, P<0.001. The distribution of unknown samples was dissimilar to that of random samples; Kolmogorov–Smirnov statistic D=0.0767, P<0.001. There were not enough data to compare the distribution of diagnostic OGTT glucose results and those of unknown provenance.

Of the test results with a recorded value, 10 504 (6.0%) were in the range for diabetes, 6661 (3.8%) were in the range for impaired fasting glucose and 245 (0.1%) were in the range for impaired glucose tolerance. A total of 6017 people (4.1%) tested had one or more glucose values in the range for diabetes when using $\geq$ 11.1 mmol/l as the diagnostic threshold for glucose results with no provenance information. By comparison, a total of 12 288 (8.4%) people had test values in the range for diabetes, when using $\geq$7.0 mmol/l as the diagnostic threshold for glucose results with no provenance information.
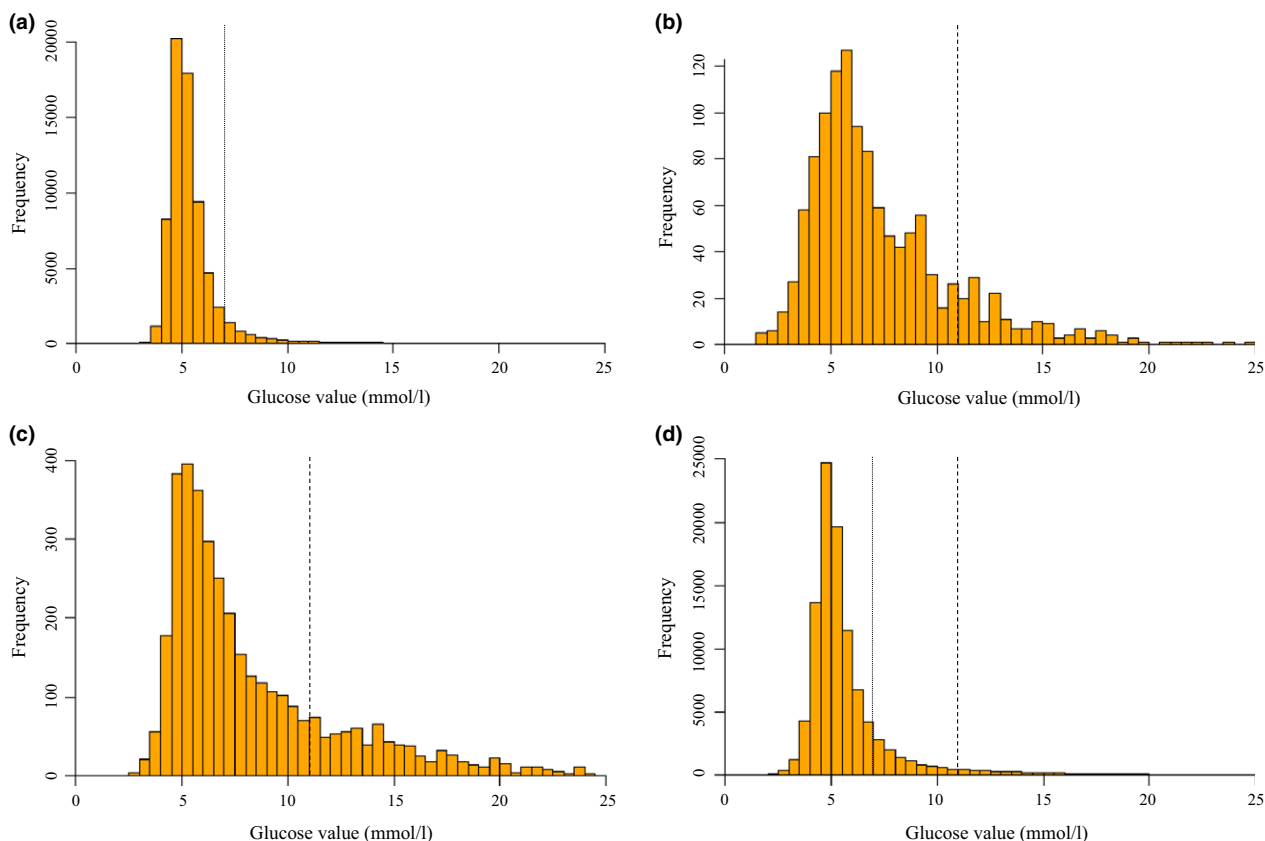
Clinical audit

A total of 460 people at the audit practice had a diagnosis of diabetes. Of these, 37 (8.0%) had Type 1 diabetes and 407 (88.5%) had Type 2 diabetes. Sixteen people had other forms of diabetes and were excluded from analysis (3.5%): one case

Table 1 Abundance of coded glucose measurements by provenance in the audit practice and the Royal College of General Practitioners Research and Surveillance Centre

Test provenance	Audit practice		RCGP RSC	
	n	% (95% CI)	*n*	% (95% CI)
Fasting glucose	68	26.6 (21.1–32.0)	75,044	36.9 (36.7–37.1)
Diagnostic OGTT glucose	1	0.4 (0.0–1.2)	1,245	0.6 (0.6–0.6)
Random glucose	24	9.4 (5.9–12.9)	6,314	3.1 (3.0–3.2)
No provenance specified	164	64.1 (58.2–69.9)	117,893	58.0 (57.8–58.2)
Other provenance specified	0	0.0 (0.0–0.8)	2,854	1.4 (1.4–1.5)
Totals	256	100.0	203,350	100.0

OGTT, oral glucose tolerance test; RCGP RSC, Royal College of General Practitioner Research and Surveillance Centre.
Glucose measurements for the audit practice are the first recorded value for each included patient.

FIGURE 1 Distribution of blood glucose values by glucose provenance: (a) fasting glucose, (b) oral glucose tolerance test glucose at two h, (c) random glucose, and (d) no provenance recorded. The dotted line indicates the diagnostic threshold for diabetes in fasted samples, the dashed line the diabetes diagnostic threshold for random samples and those taken at two h during a glucose tolerance test.

Table 2 A comparison of the distributions of glucose values by provenance

Glucose provenance	Number of samples with values	Mean (SD), mmol/l	Median, mmol/l	Skew	Kurtosis
Fasting	100 626	5.85 (2.5)	5.2	4.01	22.78
Random	3730	8.54 (4.6)	6.8	1.74	3.33
Not recorded	69 333	5.50 (1.5)	5.2	4.91	38.32

of pancreatic insufficiency, two cases of steroid-induced diabetes, three cases of gestational diabetes, three cases reclassified as diabetes resolved, four cases reclassified as prediabetes, one case where the diagnosis was never confirmed (in a temporarily registered patient) and one erroneous diagnosis. Of the included cases (*n*=444), 181 were female (40.8%). The mean (range; SD) age was 63.6 (24–99; 14.5) years.

Of those included, 16 (3.6%) patients were either diagnosed before moving to the practice or diagnosed in secondary care. From the remaining 428, the first diagnostic test was a non-OGTT glucose measurement in 256 cases (59.8%). In 122 cases it was an HbA$_{1c}$ measurement (28.5%) and in 50 cases the diagnosis was made using a formal OGTT (11.7%).

As with the RCGP RSC data, the majority of glucose tests had no provenance specified (Table 1). A total of 24 results (9.4%) had no code and therefore would be missed by an electronic search, which relies solely on pulling data using Read codes. A similar issue applies with OGTT results; only one OGTT result had an associated Read code: the remaining 49 results were identified from clinical letters, which are uploaded as attached documents. These can only be viewed manually, when searching the patient record.

Of the 164 glucose tests without provenance information, 89 (54.3%) had values between 7.0 and 11.0. These lie within a diagnostic no-mans-land.

The search identified 41 (16% of initial tests) instances where a general practitioner questioned the provenance of a glucose result, and documented this in the patient record. Of the results questioned by clinicians, 37 (90.2%) used nonspecific Read codes. The remaining four cases where provenance was questioned were when the result was coded with a fasting Read code.

Of the 41 cases where a clinician questioned the provenance of the result, 14 (34.1%) had tests repeated. One case had a substantial number of extra tests; a non-English speaking patient repeatedly had glucose tests performed; however, the clinician each time struggled to ascertain whether the sample was fasting. One repeat test was undertaken during Ramadan, and again confusion secondary to poor communication (lack of a translator) meant this also had to be repeated. This resulted in an extra five tests being performed and a significant time delay to diagnosis (614 days).

Of the 41 cases where provenance was questioned, 37 (90.2%) had a delay in diagnosis, ranging from 3 to 614 days (median 30 days; six patients had delays > 200 days).

Over time the use of Read codes at the practice appeared to change, with an increase in use of codes that include provenance information from 2009, at which time a review of glucose testing at the practice had been carried out.

Discussion

The present study highlights two major findings; the majority of blood glucose results coded (58%) in UK primary care are lacking provenance information, and this leads to unnecessary repeat tests and delayed diagnosis of diabetes.

There are implications from this research for epidemiological studies using glucose measurements drawn from routine data. Whilst HbA_{1c} results may be used increasingly for diagnosis, this is not so for retrospective studies. The distribution of results with no provenance information most closely resembles that of fasting samples, suggesting that the majority of these samples are fasting. Making assumptions about provenance for these samples would result in over- or underestimation of diabetes prevalence in the population.

The need to repeat tests because of uncertain provenance wastes doctor and phlebotomist time and is inconvenient for patients. We recognize that some of the repeat tests may be attributable to clinicians repeating the tests appropriately after 3 months to confirm the diagnosis of diabetes.

The major strength of the present investigation is that we examine the issue of failing to record glucose provenance on the both a macro and a micro scale.

The RCGP RSC covers a large population. Use of these data has provided up-to-date and nationally representative information on the recording of glucose provenance; however, the importance of this issue is not limited to the UK, with data provenance being an international issue.

The audit component of this investigation was undertaken in a single primary care centre with a special interest in diabetes research, audit and clinical coding and, therefore, may not be fully representative of the national situation. It is likely that the coding and management of diabetes in this practice is better than the national average.

Our clinical audit only included people with a confirmed diagnosis of diabetes. From this component of the investigation we are therefore unable to identify if failure in provenance recorded has led to underdiagnoses of diabetes.

Finally, we do not know what proportion of the poor data provenance originated in the practice or the laboratory. Whilst general practitioners in our audit practice expressed frustration, we could not reconstruct the data history to know at what point there was a failure to record the data required. It might have been failure by the requesting general practitioner, by the person taking the blood test or by the laboratory staff; however, it is likely to have been a combination of all three and the possibility for this failure should be build out of all elements of the system.

To the best of our knowledge no previous descriptive analysis of the provenance of blood glucose tests in a complete primary care population has been performed. We have previously reported that glucose provenance was not recorded in 65% of glucose tests performed for follow-up of women with gestational diabetes [4], which is similar to the 58% identified in the present study.

With the exception of flagging if the patient is pregnant, sufficient codes exist to enable the precise labelling of the provenance of blood glucose tests. The Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT), the most likely terminology to succeed the use of Read codes in the National Health Service and to be used internationally, also has the ability to record glucose test provenance [14].

Missing data provenance causes wasting of clinician time attempting to confirm test provenance with patients, unnecessary repeated glucose testing and delayed diagnosis. This issue could be rectified by increasing the use of existing codes that record glucose provenance data. Removing existing codes that are not used or infrequently used (Appendix S1) would improve clarity. Laboratory test request software should be set up to specify whether the value is fasted, random or part of a glucose tolerance test.

With the increasing use of routine data in healthcare research it is important to identify and manage potential data quality issues [15,16]. The provenance of blood glucose is one such challenge that will affect any research that involves defining a population with diabetes (or a related condition). Clinical ontologies offer a partial solution to data quality issues; they have been extended to provide in-depth information of the source of data [17].

The issue of the provenance of blood glucose for diabetes management and control may be less important now, with the greater use of HbA_{1c} (for which provenance is less important); however, HbA_{1c} should not be used in children or women who are currently pregnant and people with suspected Type 1 diabetes [18], and blood glucose tests with reliable provenance remain important in healthcare delivery.

In conclusion, the majority of blood glucose tests, performed in primary care, are recorded in clinical record systems without information about the provenance of the test. This causes delayed diagnosis of diabetes, unnecessary

Diabetic Medicine published by John Wiley & Sons Ltd on behalf of Diabetes UK.

5

repeated tests and wasted clinician time. Improved clinical coding in primary care is required.

References

1 World Health Organization. *Use of glycated haemoglobin (HbA1c) in diagnosis of diabetes mellitus: abbreviated report of a WHO consultation.* Geneva: WHO, 2011.

2 de Lusignan S. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. *Inform Prim Care* 2005; **13**: 65–70.

3 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.

4 McGovern AP, Butler L, Jones S, van Vlymen J, Sadek K, Munro N *et al*. Diabetes screening after gestational diabetes in England: A quantitative retrospective cohort study. *Br J Gen Pract* 2014; **64**: 22–23.

5 de Lusignan S, Liaw ST, Krause P, Curcin V, Vicente MT, Michalakidis G *et al*. Key concepts to assess the readiness of data for international research: data quality, lineage and provenance, extraction and processing errors, traceability, and curation. Contribution of the IMIA Primary Health Care Informatics Working Group. *Yearb Med Inform* 2011; **6**: 112–120.

6 Royal College of General Practitioners. Research and Surveillance Centre, 2014; Available at http://www.rcgp.org.uk/clinical-and-research/our-programmes/research-and-surveillance-centre.aspx Last accessed 21 January 2016.

7 Elliot AJ, Cross KW, Fleming DM. Seasonality and trends in the incidence and prevalence of gout in England and Wales 1994-2007. *Ann Rheum Dis* 2009; **68**: 1728–1733.

8 Hardelid P, Fleming DM, Andrews N, Barley M, Durnall H, Mangtani P *et al*. Effectiveness of trivalent and pandemic influenza vaccines in England and Wales 2008-2010: results from a cohort study in general practice. *Vaccine* 2012; **30**: 1371–1378.

9 Iturriza-Gomara M, Elliot AJ, Dockery C, Fleming DM, Gray JJ. Structured surveillance of infectious intestinal disease in pre-school children in the community: 'The Nappy Study'. *Epidemiol Infect* 2009; **137**: 922–931.

10 World Health Organization. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. 2006.

11 Sadek AR, van Vlymen J, Khunti K, de Lusignan S. Automated identification of miscoded and misclassified cases of diabetes from computer records. *Diabet Med* 2012; **29**: 410–414.

12 General Medical Council. Guidance on Good Practice: Confidentiality Guidance: Circumstances in which Patients give Implied Consent to Disclosure. 2013; (29 September 2014). Available at http://www.gmc-uk.org/guidance/ethical_guidance/confidentiality_24_35_disclosing_information_with_consent.asp Last accessed 21 January 2016.

13 National Research Ethics Service. Defining Research 2009. Available at http://nuhrise.org/wp-content/uploads/Defining-Research-Leaflet.pdf Last accessed 21 January 2016.

14 International Health Terminology Standards Development Organisation (IHTSDE). SNOMED CT The Global Language of Healthcare. 2015; Available at http://www.ihtsdo.org/snomed-ct Last accessed 21 January 2016.

15 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.

16 Powell AE, Davies HTO, Thomson RG. Using routine comparative data to assess the quality of health care: understanding and avoiding common pitfalls. *Qual Saf Health Care* 2003; **12**: 122–128.

17 Ethier JF, Curcin V, Barton A, McGilchrist MM, Bastiaens H, Andreasson A *et al*. Clinical data integration model. Core interoperability ontology for research using primary care data. *Methods Inf Med* 2015; **54**: 16–23.

18 Farmer A. Use of HbA1c in the diagnosis of diabetes. *BMJ* 2012; **345**: e7293.

Supporting Information

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

Appendix S1. Code lists used in study organised by glucose provenance. The frequencies of codes used by practices in the RCGP-RSC in 2013 are included.