A method to validate the accuracy of a centralised district diabetes register

SMR Gillani1
MRCP, MRCGP
AU Nayak1
MRCP
K Thiruvenkatasamy1
BSc, MCA
V Baskar1
FRCP, MD
BM Singh1
FRCP, MD
A Viswanath1
FRCP, MSc

1The Royal Wolverhampton NHS Trust, Wolverhampton, UK

Correspondence to: Dr Syed MR Gillani, Wolverhampton Diabetes Centre, New Cross Hospital, Wolverhampton WV10 0QP, UK; email: Syed.gillani@nhs.net

Received: 8 April 2013
Accepted in revised form: 15 May 2013

Abstract
An accurate and valid district diabetes register is needed to identify people with diabetes. Quality assurance of such a register is vital to deliver high-standard patient care. We report the findings of a methodical process of validation of the Wolverhampton District Diabetes Register (WDDR) post extraction of information from general practitioner (GP) databases, and propose an algorithm for resolving any disparity between the two data registers.

Historic diabetes register data were matched with GP databases; discrepancies were checked with GP practices and updated on the WDDR. Unidentifiable people were subject to demographic checks with the Demographic Batch Service (DBS). DBS information was used to identify patients by contacting them directly or by contacting their GP practices. Diagnostic discrepancies were corrected by biochemical checks or identifying coding errors in the GP database.

Of 2565 people unmatched with GP databases, 2380 had an identifiable GP. After checking with GP practices, 1244 (48.5%) were identified to have coding errors, 61 (2.4%) deceased and 333 (13%) with diagnostic inaccuracy of diabetes. A total of 927 (36%) patients with no identifiable GP were subject to demographic checks. Of these, 237 (9.2%) were found to be in the area and registered with another GP; 220 (8.6%) had no identifiable GP; 422 (16.4%) patients were not in the area, and 48 (1.9%) were deceased.

To maintain a valid district diabetes register (WDDR), a rolling mechanism of demographic cross checks is required at regular intervals to reduce the number of discrepancies and increase the accuracy of such a register. Copyright © 2013 John Wiley & Sons.

Practical Diabetes 2013; 30(6): 224–228

Key words
diabetes; data

Introduction
An accurate, centralised diabetes information system should identify all people with diabetes in an area and inform the delivery of a high-quality systematic, whole-population based service to include measures of access, process and outcome. The objective should be to utilise such information to minimise the adverse impact on the individual with diabetes of diabetes-related morbidity and mortality,1–3 and diminish inefficiencies in service provision – not least the resulting high expenditure on treating complications.4 Maintaining high-quality data is thus crucial to a modern-day diabetes care. A conventional approach to maintaining such registers is the extraction of information directly from general practitioner (GP) databases,5 but studies show variability in their accuracy6,7 and a number of deficiencies in quality assurance processes5 that make their validity questionable. These include poor data sharing between primary and secondary care, problems with case ascertainment, patient migration, and discrepancies resulting from multisource data input and acquisition. We are not aware of any published methodology to establish processes that test the validity of the demographic data within such primary care databases. We have therefore devised a methodical algorithm to validate all demographic information received into an established centralised diabetes register from multiple other data sources against the primary care database.

Methods
The Wolverhampton District Diabetes Register (WDDR) is a historic patient record system (dating back to 1990) for people with diabetes. It is an electronic database based on the ‘Diabeta3’ system. Since 2008, the diabetes retinal screening programme, and more recently the local foot screening programme, are all run through this database – so it is an integrated system that was...
strategically developed precisely to ensure integration of data. It links to the local hospital Patient Administration System and Pathology System and its data accuracy is checked against the national data spine. Information from individual primary care practices is separately accrued into a central data warehouse that is used as our local Primary Care Database.

To align the WDDR to the local Primary Care Database, demographic data on all alive patients coded as having diabetes from all 55 GP practices in the local area were extracted and were uploaded into the WDDR in December 2009. Individuals previously not known to the WDDR were subsequently registered. All live people registered in the WDDR who were not in the GP database were identified, and a questionnaire was sent to their GP practices to obtain information on diabetes diagnosis and demographics.

All GP practices responded to this questionnaire and information obtained was updated on the WDDR. People who were not known to have diabetes in the GP database were subject to further biochemical cross checks (previous laboratory results) to verify the diagnosis of diabetes (according to WHO criteria). The WDDR was then updated and GP practices were informed about any discrepancies for similar updating of their records. For non-identifiable patients, further demographic checks were undertaken with the Demographic Batch Service (DBS) to confirm demographic details, after which a confirmatory check with the GP practice and/or the individual if necessary was made. The data were then again updated in the WDDR.

Finally, we designed a scoring system, termed the Composite Access Score (CAS), based on three key diabetes access parameters – HbA1c, urine ACR and retinal screening – with a score of 1 being awarded for each item if done within the last 15 months. This CAS scoring was applied to the validated diabetes register to assess its utility in predicting whether an individual is in or out of area, and thus help in developing a model of validating the WDDR post extraction of information from the GP database and DBS. A score of zero would indicate that the patient has either not been active in local diabetes care or has defaulted care.

Figure 1. The results of cross checking GP databases with a central diabetes register

Live patients in WDDR (15 653) → Matched with GP database (13 088)

Unmatched with GP database (2565)

Identifiable GP in WDDR (2380)

Information from GP

Registered with GP but not DM (240)

Pathology system Biochemical checks

Diabetes (47)

IGT/GDM (67)

No DM (126)

Registered with GP but IGT/GDM (93)

Registered with GP and DM (coding error) (1244)

Deceased (61)

Not registered with any local GP (742)

DBS (927)

Both WDDR and GP databases updated

WDDR = Wolverhampton District Diabetes Register. DBS = Demographic Batch Service. IGT = impaired glucose tolerance. GDM = gestational diabetes.
Results

The results of the systematic process of validating the WDDR post data extraction from the GP database are summarised in Figures 1 and 2.

The number of live patients with diabetes from the region in the WDDR and the GP extraction were 15,653 and 13,305 respectively. A total of 217 from GP databases were not known to the WDDR and were subsequently registered.

Comparing the two registers (Figure 1), 13,088 individuals were in both databases; 2565 individuals in the WDDR were not identified in the GP database in which further processes of validation were undertaken. Of these 2565 patients in the WDDR who were not in the GP database, information on 2380 individuals was obtained by contacting their named GPs, all of whom responded. The response from GPs confirmed 1244 as registered with them and having a diagnosis of diabetes. Their data were previously not transferred by the GP practices to WDDR data transfer because of coding errors. These errors were identified, corrected and updated by locating breaks at multiple points of data transfer. There were 95 patients in GP databases who were correctly coded as either impaired glucose or gestational diabetes, and 61 patients were deceased. This information was updated in the WDDR. There were 742 individuals with no identifiable GPs.

Biochemical checks on 240 people registered in the WDDR as having diabetes, but not confirmed by their GPs, showed diabetes in 47 individuals and the GPs were informed. There was no suggestion of diabetes in 126 people, and 67 had either impaired glucose or previous gestational diabetes. These were updated in the WDDR.

Thus, an initial 185 and a further 742 known to the WDDR had no identifiable GP. Therefore, altogether, 927 individuals with no identifiable GP practice were subject to DBS checks (Figure 2). A total of 237 were confirmed to be in the region but with a different GP practice, 422 with an identifiable GP out of area, 48 deceased and 220 had no identifiable GP practice. The 457 patients who were in the area were again approached via GP and by direct patient contact and this confirmed 238 in the area, 7 deceased, and 212, with no identifiable GP, had possibly moved away.

Overall, the WDDR numbers fell from 15,653 to 14,829 (delta 824: 116 dead; 422 moved away; 286 misdiagnosed) while the GP data rose from 13,088 to 14,617 (delta 1529: 1244 miscoded; 285 not known to GPs). Ultimately, only 212 of 14,829 (1.4%) on the central register were left unaccounted for, meaning we could not identify their GP or their status by any methodology. In that small group, we applied the CAS and found 99% had CAS=0 with only two individuals with a CAS of 1, essentially confirming their inactivity in local diabetes care provision.

Figure 2. Further demographic checks among people with diabetes with no identifiable GP and not registered with local primary care
Proposed algorithm

A flow chart of this process is shown in Figure 3. We suggest that a validation process should start by doing a GP database and central diabetes register match. This match needs identification and correction of all coding errors by obtaining a complete list of codes and establishing robust data links at local database level. All discrepancies should then be subjected to demographic checks from the DBS. This will help to identify all patients who are in and out of the area. Out of area patients can be rendered inactive on a local register. Patients who are identified in the area and with an identifiable GP can be cross checked with respective GP practices. Any discrepancies can be further addressed by subjectioning them to biochemical checks. Patients who do not have an identifiable GP can be contacted directly or subject to CAS scoring to find out their current status. This will result in a validated register with greater than 98% accuracy.

Discussion

We have learned from this exercise that no one database is accurate. It is difficult to maintain the accuracy of any database due to the dynamic demographics of the population and limitations of electronic record sharing systems. The NHS Spine similarly faced several problems with its summary care record as a result of discrepancies due to technical errors in patients’ data uploads.8 However, since any such database can be very crucial to service provision to the patients, we suggest all local services should have a robust plan of validation that ensures a rolling mechanism of demographic cross checks from the DBS and GP databases at regular intervals to reduce the number of discrepancies. This identifies people with diabetes who have either moved out of the area or have moved from one GP practice to another, those who have died, those who are miscoded and those who do not have diabetes at all.

GP coding errors were found to be a major cause of discrepancy. This relates to the multiple Read Codes used by the GP clinical computer systems. The NHS Information Authority generates Read Codes regularly but GP computer system providers
A method to validate the accuracy of a centralised district diabetes register

For effective diabetes health care, accurate and reliable data are of strategic importance and, as part of data governance, data quality is paramount.

Maintenance of an accurate database is labour intensive and requires rolling mechanisms of validation from multiple database sources.

An accurate, centralised electronic diabetes register can be established by following our proposed methodology.

Key points

- For effective diabetes health care, accurate and reliable data are of strategic importance and, as part of data governance, data quality is paramount.
- Maintenance of an accurate database is labour intensive and requires rolling mechanisms of validation from multiple database sources.
- An accurate, centralised electronic diabetes register can be established by following our proposed methodology.

can also use different Read Codes for the same condition.15 In our validation process, a significant proportion of the discrepancy in our central database was found to be due to coding errors in primary care systems. There are multiple points where coding errors can happen, including completeness of data at input (GP practice level) or extraction of data by software that does not recognise all possible codes. Some patients were coded as having diabetes in GP systems but not known to the WDDR. One reason was that, on our first match, our list of codes was incomplete. A complete list of these codes can be updated at regular intervals to avoid these discrepancies. Other coding discrepancies were mainly impaired glucose tolerance (IGT) and gestational diabetes (GDM) against diabetes diagnosis, and these were all cross checked (including confirmation of the biochemical diagnosis) and corrected. Some patients were known to the register (through its triangulation with other systems [diabetes retinal screening, hospital PAS and pathology]) and, again, all of these were individually reviewed. Wherever such discrepancies were found, miscoding in the prime GP system was the most common cause of error. Completeness and correctness of data also heavily rely on the enthusiasm of practices and individual GPs.16 In our experience, there was a significant number of people with diabetes in GP practices who were not identified at our first GP contact but when, after demographic checks, we wrote back to the GP practices they managed to identify these patients on their lists.

There is a need for standardisation of Read Codes so that all people with diabetes can be identified in the electronic record system to construct and maintain a valid diabetes register since the quality of diabetes care depends on correct identification of these patients. One such example is the retinal screening programme that extracts its data from the GP2DRS system and relies on GP databases to identify people with diabetes. To achieve national standards of retinal screening, a database that is validated from multiple sources on a regular basis is vital. A recent NHS Diabetes report has concluded that 85–90% of primary care data on diabetes are fit for purpose but there is room for improvement.13 These coding errors, once rectified, will improve the quality of data fed in both the WDDR and GP2DRS databases.

The Composite Access Score is a very useful tool to identify patient access to the diabetes services. When applied to individuals who were not identifiable by GP database, DBS and direct contact, it had a strong predictive value in identifying people who have moved away. It can also be used to identify people with diabetes who are not engaged with the services. Three key elements of this score – HbA1c, urine ACR and retinal screening – will incorporate all the process modalities (blood test, urine test and contact with screening programme) a patient should have as part of their routine diabetes care. The National Diabetes Inpatient Audit (NaDIA 2010–2011) report suggested that failure to complete nine care processes could deprive people of timely intervention to prevent complications.12 By applying a CAS score to all patients on a district diabetes register, all non-engaged people can be identified and a proactive approach can be adapted to get these patients involved in their diabetes care.

Biochemical checks could be utilised effectively to validate a database and to populate relevant registers such as the WDDR, IGT and GDM registers. It will help to identify accurately the people to whom to offer correct management advice according to their status of diabetes or pre-diabetes. This fail-safe process minimises the chances of inadvertently removing an individual from the register purely based on information available from GP data extraction that may not be always accurate. This whole process of data validation is tedious and labour intensive, albeit worthwhile. Once a valid database is established after going through all of this hard work, it is easy to maintain it by cross checking details with other databases at regular intervals. The number of discrepancies on following checks has significantly reduced, making it a quick and cost effective process. It has not only improved completeness of key care processes including retinal and foot screening programmes, but has also helped us to plan our service delivery and resources according to the needs of the population. After this whole process of validation we now have a >97% accurate district diabetes register in Wolverhampton.

Declaration of interests

There are no conflicts of interest declared.

References