ROBERTSON, A.R.R., FERNANDO, B., MORRISON, Z., KALRA, D. and SHEIKH, A. 2015. Structuring and coding in health care records: a qualitative analysis using diabetes as a case study. *Journal of innovation in health informatics* [online], 22(2), pages 275-283. Available from: https://doi.org/10.14236/jhi.v22i2.90

Structuring and coding in health care records: a qualitative analysis using diabetes as a case study.

ROBERTSON, A.R.R., FERNANDO, B., MORRISON, Z., KALRA, D. and SHEIKH, A.

2015

JOURNAL OF
INNOVATION IN
HEALTH INFORMATICS

Research article

Structuring and coding in health care records: a qualitative analysis using diabetes as a case study

Ann R R Robertson
eHealth Research Group, Centre for Population Health Sciences,
The University of Edinburgh, Edinburgh, UK

Bernard Fernando
eHealth Research Group, Centre for Population Health Sciences,
The University of Edinburgh, Edinburgh, UK

Zoe Morrison
eHealth Research Group, Centre for Population Health Sciences,
The University of Edinburgh, Edinburgh, UK

Dipak Kalra
Centre for Health Informatics and Multiprofessional Education,
University College London, London, UK

Aziz Sheikh
eHealth Research Group, Centre for Population Health Sciences,
The University of Edinburgh, Edinburgh, UK
Division of General Internal Medicine and Primary Care,
Brigham and Women's Hospital/Harvard Medical School,
Boston, MA, USA

Cite this article: Robertson ARR, Fernando B, Morrison Z, Kalra D, Sheikh A. Structuring and coding in health care records: a qualitative analysis using diabetes as a case study. *J Innov Health Inform*. 2015;22(2):275–283.

http://dx.doi.org/10.14236/jhi.v22i2.90

Author address for correspondence:
Ann R R Robertson
Centre for Population Health Sciences
Teviot Place
The University of Edinburgh
Edinburgh EH8 9AG, UK
Email: A.R.R.Robertson@ed.ac.uk

Accepted October 2014

ABSTRACT

Background Globally, diabetes mellitus presents a substantial and increasing burden to individuals, health care systems and society. Structuring and coding of information in the electronic health record underpin attempts to improve sharing and searching for information. Digital records for those with long-term conditions are expected to bring direct and secondary uses benefits, and potentially to support patient self-management.

Aims and objectives We sought to investigate if how and why records for adults with diabetes were structured and coded and to explore a range of UK stakeholders' perceptions of current practice in the National Health Service.

Methods We carried out a qualitative, theoretically informed case study of documenting health care information for diabetes in family practice and hospital settings in England, using semi-structured interviews, observations, systems demonstrations and documentary data.

Results We conducted 22 interviews and four on-site observations. With respect to secondary uses – research, audit, public health and service planning – interviewees clearly articulated the benefits of highly structured and coded diabetes data and it was believed that benefits would expand through linkage to other datasets. Direct, more marginal, clinical benefits in terms of managing and monitoring diabetes and perhaps encouraging patient self-management were also reported. We observed marked differences in levels of record structuring and/or coding between family practices, where it was high, and the hospital. We found little evidence that structured and coded data were being exploited to improve information sharing between care settings.

Conclusions Using high levels of data structuring and coding in records for diabetes patients has the potential to be exploited more fully, and lessons might be learned from successful developments elsewhere in the UK. A first step would be for hospitals to attain levels of health information technology infrastructure and systems use commensurate with family practices.

Keywords: clinical coding, diabetes mellitus, medical records, qualitative research

INTRODUCTION

The number of people worldwide with the long-term, complex endocrine condition of diabetes is predicted to rise above 470 million by 2030, linked to ageing populations, ethnicity and lifestyle-related factors, such as obesity.[1–4] Type 2 diabetes is a complex and potentially serious long-term condition, which is costly in terms of health care resources and patient morbidity and mortality. While there are individual differences in its presentation and prognosis that are not yet entirely understood, the disease and its complications are largely preventable, or controllable, with active management by patients themselves and by their health care professionals.[4] Hence, as with other long-term health conditions, the effective management of diabetes is a high priority for health service providers.[5]

In the UK, the National Health Service (NHS) has developed a multi-faceted approach to managing diabetes. The approach encompasses: improving the early detection, diagnosis and monitoring of diabetes; more timely interventions that can often involve a range of health care specialists in addition to the family doctor and supporting patient self-management and planned, structured service provision.[3] The activities within this approach rely on good quality, accurate information being documented, accessible and sharable. To achieve integrated care, patient information needs to be available – and comprehensible – to a range of health and social care staff, over geographical distances and over time.[6]

Currently, structured and coded digital information underpins efforts to achieve integrated health care delivery and other benefits. Developing a national approach to more effectively manage diabetes therefore aligns with implementing health information technology (HIT) systems and electronic health records (EHRs) in family practices and in hospitals. Nonetheless, an incorrectly coded or incomplete record could adversely affect clinical management decisions and also undermine research findings based on (inaccurate) routinely collected data. Furthermore, there might be a clinical need to record information that is not amenable to structuring or coding, hence valuable information may potentially be lost if structured and coded data capture is overly emphasised.

As part of a larger research programme that aimed to inform best practice in relation to optimum levels of structured and/or coded data in NHS EHRs,[7,8] the aims of this case study were to: investigate if how and why records for adults with diabetes were being structured and coded in the NHS in England and explore a range of UK stakeholders' perceptions of documentation approaches for patients with diabetes.

METHODS

Design

We undertook a qualitative case study.[9] The 'case' was the documenting in England of diabetes health care provision in patients' records, that is if and how information on these patients is being digitally captured and transmitted. An in-depth case study allows the description of any currently used technologies (hardware and software) and work practices, with a focus on how and why different forms of data capture – free-text, narrative, structured data (e.g. using a template) and coded data (e.g. with Read Codes) – are being used in these health records. The case study design also allows a detailed exploration of a range of views on the optimal balance between unstructured, structured-uncoded and structured-coded data recording, particularly for clinical care but also for NHS management, diabetic service planning, research and audit. We sought to illuminate how the ways in which data were documented facilitated the management of the long-term condition, for example through call and recall prompts, health status monitoring and supporting communication between health care professionals.

Sampling

Two primary care sites and an NHS hospital in the same region in England were identified as possible case study participants through the UK Clinical Research Network.[10] Choosing associated family practice and hospital sites was designed to give a coherent picture of a geographical health community, in which the majority of patients diagnosed with diabetes receive care. It also allowed us to generate context-specific findings from the sites as well more generic findings that were likely to be generalisable from considering the sampled sites together as a holistic 'case'.

We used purposive sampling[11] within the sites and beyond them to identify and recruit a wide range of individuals who affect or are affected by the processes and outcomes of documenting diabetes patients' health care. Interviewees outside the sites were selected to provide contextual information for the case study and to provide additional perspectives. Throughout, we actively sought to sample a range of experiences and views, including outlier views. Interviewees included clinical staff (doctors, nurses and allied health professionals); patients; NHS managers; clinical coders; clinical systems developers/suppliers and academics.

Data collection

Initially, we purposively recruited and interviewed eight individuals in England and Scotland with a particular interest in diabetes and structuring and/or coding of health records in order to gain their perspectives as academics (including academic clinicians) and health care informatics experts (including developers of HIT systems). This was achieved through research team contacts and snowballing, whereby one informant suggests another informant. These early interviews allowed the subsequent interviews in the case study sites to be conducted in the context of information about the situation more widely in the UK and beyond. Interviews were guided by initial topic guide schedules derived from the research protocol, which were adapted by the case study's lead researcher as the investigation progressed. Audio files of recorded interviews were transcribed professionally. Additional data were collected in field notes taken during two non-participant observations in a family practice diabetes review clinic and naturalistic, on-site observations in hospital. Although this study did not aim to assess the technologies, demonstrations of some HIT systems commonly used for diabetes management in the UK provided further background information. We also reviewed relevant documentary data, including diabetes-related websites, documentation on different coding systems and current local and national diabetes-related guidelines and protocols, as background information before undertaking the case study interviews and observations.

Data analysis

Thematic data analysis[12] and data collection were iterative, allowing identified themes to be further explored and any discrepancies between individuals' accounts or between data sources to be further investigated. Interview transcripts and field notes were uploaded into the qualitative software package NVivo 8.[13] For the thematic analysis, coding was guided by an initial framework of categories derived from the research protocol, a literature search and research team discussions. These categories were refined as the interview data collection progressed. Coding combined top-down coding, based on the coding framework, and bottom-up coding in which new or more fine-grained themes emerge out of the data.[14] Interpretation of the findings was influenced by a socio-technical theoretical understanding, i.e. that technologies and those who work with them function inter-dependently.[15] Approaches to validating data quality and credibility in this qualitative case study included checking data for face validity, looking for disconfirming evidence, data triangulation by data source and discussing data collection and analysis in the wider team throughout the research process.[14,16]

Ethical considerations

We obtained Research Ethics Committee approval and local site permissions. Interviewees gave informed consent before being recruited, and all data were anonymised to protect participant confidentiality.

RESULTS

The case study dataset reported here comprised 22 semi-structured interviews and field notes from four observations in two family practices (n = 2) and one hospital (n = 2). The occupations of the interviewees are given in Table 1. The main themes derived from the interview data related to the different contexts of the family practice and hospital care settings ('Working in different worlds'); the limited electronic exchange of information between the two sectors ('Communicating across the divide?') and the different approaches to the structuring and coding of diabetes data we observed ('Drivers; Who codes and how?; Involving the patients and Coding enablers and barriers'). The final theme reported here related to study participants' views of the consequences of coding ('Clinical and research benefits'). The initial coding categories by which the data were organised and the seven themes derived from these are given in Table 2.

Table 1 **The interview dataset for this case study**

interviews n = 22
Contextual – 7
(6 academics/clinicians and 1 system supplier)
Family practice staff – 5
(1 family doctor; 2 diabetes nurses; 1 practice manager and 1 coder)
Diabetes patients – 2
Hospital staff – 8
(1 diabetes consultant; 1 junior doctor; 1 ward nurse; 1 diabetes research nurse; 1 dietician; 2 coders and 1 data manager)

Table 2 **Initial coding categories and extracted themes**

Category	Amounts and types of structured and/or coded information for diabetes care
Themes	**Working in different worlds** and **Communicating across the divide?**
Category	Approaches to coding for diabetes
Themes	**Drivers for structuring and/or coding; Who codes and how?; Involving the patients and Coding enablers and barriers.**
Category	Perceived consequences of more coding
Themes	**Clinical and research benefits.**

Working in different worlds

Our case study illustrated a marked contrast between the sampled sites. In a busy NHS teaching hospital, diabetes patients admitted into hospital wards were typically very unwell, for example with kidney or heart failure, or they were maternity patients. The hospital consultants and their teams looked after hospital inpatients. Hospital consultants also saw diabetes patients in community clinics. Here, they worked in collaboration with community nurses, although not with family practice doctors. Hospital doctors relied on paper to record information about diabetes patients. The hospital had implemented a number of different HIT systems, such as a pathology system, but had no comprehensive EHR system.

'There are plans to introduce EHRs in the future. The hospital X-ray system is different from the system for blood results, which is different from the patient management system. At the moment, I have to log into about at least five different systems to get information on a patient. So I created my own diabetes template. Yes, it's a paper template and I use it for my patients, and a couple of other consultants have been using it for their clinics too'. (Hospital Consultant)

In contrast, the context in which family practice staff worked was a self-contained, computerised surgery, where the use of HIT systems and coding were well established.

'We've got a good system in place. We've been doing it a long time here. We code everything. Whoever is entering it does a Read Code'. (Family Practice Doctor)

There were four types of data capture using the family practice HIT system's diabetes template: numerical values (e.g. weight and height); yes/no options (e.g. smoking); dropdown menus (e.g. diagnosis) and free text boxes (e.g. for comments on dietary advice that had been given). In addition, information about diabetes patients and their care was recorded in free text in a separate part of the EHR, which was also coded. There was minimal free text that was unrelated to an entered code. For example, a family doctor described coding all her entries in a patient's record with the exception of a free text comment on her reasons for altering the patient's management plan. Unlike in the hospital setting, where the paper template meant handwritten entries were structured but there was no coding in the record, the family practice record entries were predominantly coded.

Communicating across the divide?

The family practice was the main health care setting in which the majority of diabetes patients received most of their care. Nonetheless, some patients would be referred to community clinics to be seen there by hospital consultants, and there were some patients who had episodes of inpatient hospital care. NHS staff interviewees from each sector described largely working in isolation from the other. Letters were the most common method of communicating information between the hospital and family doctors. Phone and fax messages might be used instead of paper if the message was more urgent. Digital data sharing across the care boundary was restricted to some laboratory results, which would then be coded in the family practice HIT system.

'Everyone functions sort of independently. We get each other's letters. Most letters are probably accurate, but it's not in real time'. (Hospital Doctor)

From a system supplier's perspective, the difficulties of modernising communications between family practices and hospitals with HIT systems were less technical than political and economic.

'We can put what you like in the can, but there are very few good examples of good integration between primary and secondary [hospital] care in England. The reason is partly from a lack of will – there is demand, lots – but no-one is picking up that ball. How to fund it? Also, there are some issues around information governance and so on. GPs [family doctors] are cautious about who sees their data and so on. The majority of processes are not integrated across primary and secondary care at all'. (Systems Supplier)

A hospital doctor described dictating his letters to the family practice in front of the patient, and a copy of the letter being returned to him for review and signing after it had been transcribed by a secretary, then the letter being posted. This process took several days, with hospital test results often following later. Electronic document transfer would be a relatively straightforward step that would reduce some delays in passing patient information from one sector to the other and allow patients' records to be updated in a timely manner:

'For communications, it's normally in the form of a letter, and there are tools for configuring and producing letters, and it's possible to convert to electronic document sharing at some sites - but integration – that's the next leap'. (Systems Supplier)

Elsewhere in the UK, the integration gap for diabetes patients was being addressed through HIT. In Scotland, a national, computerised system, now called SCI-Diabetes, still widely known as SCI-DC,[17] combined information from a number of sources, including from podiatrists, dieticians, hospitals and the family doctor, to give diabetes patients a single, quickly updated record.

This clinical system had taken some 10 years of investment to collect the data and set up the systems, building out from a smaller regional system. Interviewees in Scotland said the development of SCI-DC had benefitted from the drive and enthusiasm of particular individuals who had been its 'champions' but they were unanimous in reporting that the long effort had been sustained by the interest and support of the clinical community more widely.

'The approach was always to work with those who wanted to sign up to the system and let the uninterested or refusers see the benefits over time and join in later. Now nearly all GPs [family doctors] and all hospitals are signed up… It's about integrating the whole thing into work patterns, and also about the rewards for clinicians, such as having blood results available for the consultation, not having to dictate results anymore and having an organised letter. Scotland has gone a wee bit independent, and is doing very well'. (Hospital Doctor/Academic)

Drivers for structuring/coding diabetes information

The sampled sites shared a financial incentive to code diabetes data. For the hospital, it was income derived from reporting activity under the payments by results (PbR) scheme. PbR payments accounted for a significant proportion of the hospital's overall income. All participants at the sampled hospital site reported they were conscious of the financial difficulties facing the NHS during a period in which employees' posts were being reviewed and the clinical workforce was being pared back. However, not all made the link between coding, PbR payments and the hospital's finances.

'I don't think I ever thought about it until I started research work. You go all through your training and so on, and coding and its importance is never mentioned, or if it is, nobody takes it in. It's just something that goes on behind the scenes. You don't even know where the hospital coders are'. (Hospital Research Nurse)

Similarly, family practice interviewees mentioned payment-related quality outcomes framework (QOF) points as being one driver for coding.

'It makes sense if everything is coded correctly because it is income for the practice'. (Family Doctor)

In addition to the financial driver, family practice staff referred to a 'culture' of using computers and coding in doctors' surgeries, where computer systems had been in use for years. Clinical and non-clinical family practice staff recognised the need for coding in order to be able to carry out searches and audits as part of their everyday work. They were also motivated to code because having coded information was perceived to improve direct patient care.

'Codes mean you can search and identify patients, say, who need six monthly blood pressure checks or who have not turned up for a B12 injection. It does improve patient care'. (Family Practice Coder)

Improving patient care was also given as the reason for structuring patient information using the hospital consultant's paper templates. Paper templates were a pragmatic solution to encourage standardised, clinical practice and to aid the completeness of recording patient information in the absence of hospital EHRs and digital templates.

'In an ideal world, yes, I would like a diabetes HIT system but in real life, if I aim for it, it won't happen because of finance. So I just got on with this. From the first day someone hits my ward, they get a yellow sheet and all the co-morbidities, medications, interventions are there, and we update the yellow sheet on a regular basis. It forms the basis of the discharge letter. Simple things, like pre-conception counselling – it is very easy to miss it otherwise'. (Hospital Consultant)

Who codes and how?

We observed marked differences between settings with respect to who did the coding and how this was done. Read codes were used in family practice. This is an extensive list of terms, each attached to a unique code and which covers information ranging from the patient's occupation to signs and symptoms, tests, diagnoses, interventions to medications and more. In the family practice surgeries, it was predominantly diabetes specialist nurses who completed digital diabetes templates during, or sometimes shortly after, diabetes review appointments. Read coding was embedded in the digital template. The family doctors were more likely to use another part of the EHR for their consultations with patients with diabetes, and to code those entries themselves. Another staff member also had a specific coding role, whether the job title was 'coder', 'clerk' or 'secretary'. This individual would complete data entry and coding in the EHR when information about a patient arrived from the local hospital.

'Filling in the template automatically fires codes behind the scenes. It is embedded in the system. We always code for type 1 or type 2 diabetes. I would say nine out of ten templates are fully completed. Obviously, some of it can't be done during the review appointment, for example when someone has an eye check at the opticians, obviously we don't do that here, but when the paperwork comes through from the optician, or the hospital letters with, say, blood pressure results, I put it on the patient's template. All letters normally get sorted within a couple of weeks. If there was anything urgent, they would fax, but normally it's routine results by letter and they get updated and coded in the patient's record within a couple of weeks'. (Family Practice Coder)

In the hospital, coders used International Classification of Disease (ICD-10) codes for diagnoses and Office of Population, Censuses and Surveys Classification of Surgical Operations and Procedures (OPCS-4) codes. Coding was based on the content of patients' discharge summaries which were sent to a team of professionally trained, full-time coders. We observed the coding team accommodated in one large room in the hospital, where the coders worked from their own copies of coding manuals that were extensively annotated and individualised with comments, reminders and coding updates. They dealt with coding queries arising from information that was unclear, inconsistent or missing from the discharge sheet by telephoning the doctor who had written it. An interviewee described how they were tightly constrained by the content and completeness of the discharge summaries that doctors produced. There were variations between hospital departments and between individual doctors, which professional coders came to recognise:

'We have one doctor who doesn't like calling people 'geriatrics' so she won't put down a geriatric fall, she will put down a mechanical fall, but there is no code for mechanical fall – so we have to phone her up. And it could be that the person fell and broke their leg because they were hypo and really it was diabetes that caused the fall but that won't be what's on the discharge summary, it'll be surgery for a broken leg, so that's what's coded'. (Hospital Coder)

Coding based on discharge summaries, rather than the whole record, had implications beyond missing diagnoses codes. Interviewees discussed how some activities also became 'invisible'. For example, a dietician would spend time on a non-diabetes ward giving dietary advice to a patient with diabetes, but it would be the activities relating to the primary reason for patient's admission – not diabetes – that would feature on the discharge summary and hence get coded.

In some instances, a coder might work in a clinical department and code directly from the patient's hospital record. This depended in part on the availability of accommodation for the coder in the department and on the coding interest of key individuals in that speciality. A coding manager reported this was time consuming but resulted in more complete and accurate coding. The hospital's coding department was being relocated and losing a number of desks, which meant that in future coders would be more likely to do at least some of their work from home.

'Coding from the patient's record has real advantages over coding from discharge summaries. That would be the ideal. A big part of what we are trying to do now is to get out there, for coders to meet clinicians, so at least they remember a face, and they understand what we are doing and why it is important. But we are juggling with other considerations too, like fewer desks for us in the new hospital'. (Hospital Coder)

Another participant described noticing misclassification of diabetes type in the course of research work she had conducted at the hospital, based on her own specialist knowledge of the disease.

'I could see type 1 and knew it was not but it got put on because of insulin and once it was coded that way…' (Hospital Research Nurse)

Similarly, a family doctor remarked on the problem of miscoding of type 1 and type 2 diabetes in the family practice HIT system, particularly from coding in the past:

'The only problem is where someone is coded as insulin dependent, and is it truly type 1 or type 2? Some of the old codes in the system need a bit of tidying up'. (Family Doctor)

In Scotland, the coding approach for the national research register derived from the SCI-DC system[17] included inferring diagnoses of diabetes. Recognising that statistics from any one source were unreliable, such as a diagnosis of diabetes being missed from a hospital discharge summary, SCI-DC was designed to reconcile diabetes data from multiple sources, including laboratory data, the retinopathy screening programme, podiatry and data from both secondary and primary care, in order to improve the quality of the aggregated dataset.

'We would like all diabetes diagnoses and check-ups to be properly coded but this is not done. We are conservative in making any inferences, we err on the side of caution. For example a blood sugar check, no, but if a GP had coded for a diabetes clinic check-up, although they have not coded a diagnosis of diabetes, yes, we could infer a diagnosis of diabetes'. (Clinician, NHS National Services Scotland)

Involving the patients

Family practice nurses were observed using the HIT system to generate graphs on screen to illustrate trends for that individual in such measures as weight and glucose control. The graphic illustration was accompanied by the nurse's reassurance and praise for good control or was used to initiate a discussion with the patient about management where the trend showed some deterioration. Both nurses and patients reported finding this helpful. Echoing the hospital consultant who used a paper template to ensure the completeness of recorded information, nurses said the digital template was useful for making sure nothing was missed during the appointment. Having items on the screen prevented 'skipping bits' or forgetting to go back to an uncompleted section before the close of the consultation.

Participants perceived that accessing the HIT system during diabetes review appointments facilitated the nurse–patient interaction, or was neutral. It allowed quick comparisons of new test results with previous results during the consultation, and was thought potentially useful for encouraging patients' involvement in their own care. Using the HIT system well was seen as part of the health professional's consulting skills.

'Part of the skill of using it is that it isn't a problem. I suppose there could be exceptions, patients who didn't like having a computer in the room – but I honestly can't think of an occasion. I can't think of anyone ever picking me up on it!' (Family Practice Nurse)

Barriers to and enablers of coding

Hospital coders sampled in this case study were comfortable that they had the necessary training, support and competence to code. They were supported to undertake additional professional qualifications and took pride in their medical knowledge and coding skills. There was evidence of job satisfaction, partly derived from a belief in the importance of what they were doing. The principal barrier they perceived to better coding was the limitations of working from hospital discharge summaries.

Similarly, coders in family practice described good access to support and training, and in their case also having an HIT system that was by and large easy to use because much of the coding was automated.

Being part of a small team in a family practice setting was also seen as helpful:

'We all know each other and get on. We sit in the same room as each other so if there's a query or a problem, or you notice something is missing that needs coding, you just ask the person'. (Family Practice Nurse)

Interviewees were asked their views on a government proposal for all NHS organisations in England to change to using a new clinical terminology. The commercial supplier participant was well aware of a proposal to have SNOMED-CT codes used by the whole of the NHS in England in future; however, no other interviewee at the time of data collection knew about SNOMED-CT codes or when any change to using these codes might happen.

'We haven't got anybody in the NHS asking us for it, not yet, but we are fully aware and know what to do'. (Systems Supplier)

'SNOMED? What's that? I haven't heard of that. … It's not on our radar yet!' (Family Practice Coder)

None of the family practice interviewees indicated that difficulties with using their HIT system posed a barrier to coding for them. One nurse reported there were some 'niggles', such as having to click to change screens, but in her experience these were not a problem:

'No, I'm positive about it'. (Diabetes Nurse)

Clinical and research benefits

Both family practice and hospital clinicians in our study identified some clinical benefits from structuring and coding diabetes data. These related to the availability of information for consultations, disease management by protocol and speed of information sharing in the case of electronic transfers of EHRs for patients moving from one family practice to another. A far greater potential benefit, according to some interviewees, lay in secondary uses of coded data. Clinical interviewees and academic clinicians believed that coding would lead to significant improvements in research, public health, epidemiology, service planning and audit. These participants were based in the case study settings and in academic institutions outside those sites.

'At the moment, I can't answer even basic questions about our diabetes patients, unless I go through all the records one by one and can you imagine how long that would take? Coding is fundamental for researchers'. (Hospital Research Nurse)

There was particular enthusiasm for the SCI-DC system in Scotland, which supported regional diabetes registers, prompting screening calls, and research participation through a national patient research register.

'SCI-DC is the way ahead. Everybody is going to find their way on to it somehow because it uses multiple sources. We can extract more value from SCI-DC by linking. In effect, we have potentially huge cohort studies. There is enormous potential value through linkage for public health'. (Hospital Doctor/Academic)

In addition to cohort studies, interviewees suggested that the availability of a comprehensive dataset of the population with diabetes would lead to wider public participation in medical research and held the promise of easier, faster and cheaper, large-scale randomised controlled trials.

'The diabetes dataset we have in Scotland is among the best in the world, and researchers from other countries are coming to us now wanting to use it too. It links to other datasets, such as the cancer register and maternal data, which makes it hugely valuable for future research'. (University Researcher)

DISCUSSION

Our case study showed high levels of diabetes-related data structuring and coding in family practices. However, the levels of data structuring and/or coding in diabetes records were seen to vary considerably between family practices and hospitals, and between hospitals. We found that the advantages of highly coded diabetes data for secondary uses, such as research, public health and service planning – especially if diabetes data can be combined with other disease registers, as in Scotland – were considered by interviewees to be potentially immense. For example, interviewees reported they held the promise of revolutionising medical research by making randomised controlled trials quicker and cheaper to run and by encouraging large cohort studies through supporting assessment of eligibility criteria and enabling more targeted recruitment decisions.

We also found interviewees reported more marginal, direct clinical benefits within an organisation. Perceived clinical benefits included: having relevant, up-to-date information available for consultations; making it easier to share information between members of staff; improving disease management and monitoring; helping to involve patients in their care by illustrating points graphically on the screen and enabling fast transfer of up-to-date records when patients moved practice. This study found no evidence in the sampled sites of other hoped-for benefits, such as improving communications between different NHS organisations and supporting better integrated care, or of enhancing patients' self-management. Nonetheless, there was the SCI-DC[17] model developed in Scotland, where a single, diabetes patient record was created from multiple information sources, which included the family doctor and the hospital.

Case study interviewees using family practice HIT systems were in the main satisfied with their systems and felt data coding was relatively straightforward and worthwhile. Our sampled hospital site used paper records. Clinicians spoke in favour of having a specialist diabetes HIT system to support more structured data capture, but these interviewees reported seeing no prospect of getting such a system in the foreseeable future because of financial constraints in the NHS. There was widespread awareness of problems arising from inaccurate and incomplete coding, among clinicians, managers and coders. There was virtually no awareness among NHS staff in the case study sites of any plans to make SNOMED-CT codes the universal standard in the NHS in England.

Our data supported previous reports of the misclassification of diabetes type 1 and type 2 and wider concerns about inaccurate and incomplete coding of diabetes data, which leads to the distortion of diabetes-related statistics.[18–23] In addition, our detailed case study of a sample health community illustrated the wide discrepancies in the amounts, types and approaches to coding that can currently be found across health care sectors in England. A common factor was that financial incentives were a driver for coding, but it was one amongst other drivers for coding for our interviewees who also discerned some direct, clinical benefits. Furthermore, in our sampled hospital, the reported importance of coding for financial reasons varied by role, with coders and managers more aware of the implications of coding for the hospital's finances than clinical staff. Rather, in our work, hospital clinicians viewed the hospital's finances to be the main barrier to implementing specialist diabetes HIT systems and hence more structured and/or coded data.

While previous research has investigated the influence of EHRs on the quality of the management of diabetes patients,[24,25] our study suggested a widespread perception in family practice that structured and coded information had direct benefits for clinical care. The observed emphasis on highly structured and coded data capture in family practice did not seem to discourage recording of any aspects of the clinical consultation, as has sometimes been speculated.

Strengths and limitations

The strengths of this case study are the range of interviewees who described their experiences and views on data structuring and coding for diabetes and the sampling of three sites from one area (two family practices and one hospital), which gave a picture of diabetes information recording across a local health community. We selected diabetes as an exemplar of long-term, biomedical conditions. As such, our diabetes-related findings may at least some extent be transferrable to other long-term conditions, for instance to asthma, chronic obstructive pulmonary disease and stroke, in which on-going monitoring and disease management by health professionals and by patients are also important factors in care.

The main limitation of the work also relates to the highly variable approaches to diabetes data capture in English hospitals. Our detailed, on-site findings are restricted to one health community in which the sampled hospital used paper-based diabetes records; had resources permitted, it would have been valuable also to sample a hospital in which EHRs were established. Qualitative research elicits the subjective accounts of participants. While this paper provides an accurate account of the interview data we collected, we acknowledge that data saturation was not necessarily achieved. There may be quite different experiences and perspectives that we failed to capture here, despite interviewing a diverse range of clinical and non-clinical participants.

CONCLUSIONS

Examples from regions in the UK where high levels of diabetes data structuring and coding are in common use – such as most family practices in England and in the smaller UK nation of Scotland – indicate the value of striving to have this approach adopted more widely. There is potential for substantial secondary uses benefits, notably for research and for more marginal clinical benefits within individual health care organisations. Where NHS staff already work well with existing clinical codes, they will have to be persuaded of the benefit of replacing familiar codes with a new clinical terminology should the government revive plans to roll out SNOMED-CT.

Scotland has a significantly smaller population than England. Differences in scale and in the national political environment mean developments in Scotland are unlikely to be able to be directly transplanted to England or other large nations. Nonetheless, Scotland offers an example of successful progress and innovation; the model there of slowly rolling out a largely standardised approach to digital diabetes data collection, which was always clinically led, and of combining multiple data sources to improve the quality of secondary uses data, rather than focusing on improving individuals' coding, is one which other countries in the UK and further afield could usefully study and perhaps follow.

Initially, efforts in England might be best directed towards reducing the discrepancies between family practices and hospitals with respect to HIT, and particularly the variability among hospitals in how diabetes data are currently captured, stored and shared. If specialist diabetes HIT systems became widely established in NHS hospitals in England, the next steps could then focus on using digital data for improving real-time information sharing between the different healthcare sectors and between those sectors and community health and social care providers. Coding from hospital EHRs would be preferable to coding from hospital discharge letters, and digital documentation could be further exploited to support patients' access to information about themselves and how to manage living with diabetes.

Contributorship

AR, BF, ZM, DK and AS conceived and developed this case study and participated in analysing the data and writing the report. AR collected the data and drafted the case study report. AS was the principal investigator.

Acknowledgements

We are grateful to the members of our Independent Project Steering Committee, Professor Simon de Lusignan (chairman), Dr Nick Booth, Dr Stephen Kay and Lee Priest, and to all the NHS staff and interviewees who gave generously of their time in support of this research.

Funding

This research was funded by the NHS CFH Evaluation Programme (NHS CFHEP 009). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NHS CFH Evaluation Programme or the Department of Health. AS is supported by The Commonwealth Fund.

Competing interests

None

REFERENCES

1. The Lancet editorial. The diabetes pandemic. *The Lancet* 2011 (July);378:99.
2. Royal College of General Practitioners & NHS Diabetes. Coding, classification and diagnosis of diabetes report. 2011 (March). Available at: http://www.sccgformulary.co.uk/Coding%20Classification%20and%20Diagnosis%20of%20Diabetes%20Report.pdf.
3. Department of Health. Six years on: delivering the Diabetes National Service Framework. 2010(Feb). Available at: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_112511.pdf.
4. Department of Health. Management of adult diabetes services in the NHS 2012 (May). Available at: http://www.nao.org.uk/wp-content/uploads/2012/05/121321.pdf.
5. Department of Health. Your health, your way – your NHS guide to long-term conditions and self care. Available at: http://www.nhs.uk/Planners/Yourhealth/Pages/Yourhealth.aspx. [last accessed 03/03/2014].
6. Department of Health. Health and Social Care Bill. 2011. Available at: http://www.publications.parliament.uk/pa/cm201011/cmbills/132/11132.pdf.
7. Morrison Z, Fernando B, Kalra D, Cresswell K and Sheikh A. National evaluation of the benefits and risks of greater

structuring and coding of the electronic health record: exploratory qualitative study. *JAMIA*. doi:10.1136/amiajnl- 2013-001666. Available at: http://jamia.bmj.com/content/early/2013/11/01/amiajnl-2013-001666.full.

8. Morrison Z, Fernando B, Kalra D, Cresswell K, Robertson A and Sheikh A. The collection and utilisation of patient ethnicity data in general practices and hospitals in the United Kingdom: a qualitative case study. *Informatics in Primary Care* 2014;21(3):118–31. PMid:25207615.

9. Stake RE. Case studies. Denzin NK and Lincoln YS. (Eds). *Handbook of Qualitative Research*. London: Sage Publications, 1994, pp. 236–47.

10. UK Clinical Research Network. Information available at: http://www.ukcrc.org/research-infrastructure/clinical-research-networks/uk-clinical-research-network-ukcrn. [last accessed 09/11/2013].

11. Patton MQ. *Qualitative Evaluation and Research Methods*, 2nd edition. London: Sage Publications, 1990.

12. Braun V and Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006;3(2):77–101. Available at: http://dx.doi.org/10.1191/1478088706qp063oa.

13. 12NVivo8. QSR International. Information available at: http://www.qsrinternational.com/products_previous-products_nvivo8.aspx. [last accessed 09/12/2013].

14. Corbin J and Strauss A (Eds.). Strategies for qualitative data analysis. *Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory,* 3rd edition. London: Sage Publications, 2008.

15. Coiera E. Four rules for the reinvention of health care. *British Medical Journal* 2004;328:1197–9. Available at: http://dx.doi.org/10.1136/bmj.328.7449.1197. PMid:15142933; PMCid:PMC411109.

16. Silverman D. *Doing Qualitative Research: A Practical Handbook,* 1st edition. London: Sage Publications, 2000.

17. The Scottish Care Information – Diabetes Collaboration (SCI-DC). Information available at: http://www.sci-diabetes.scot.nhs.uk [last accessed 09/12/2013].

18. Farmer A and Fox R. Diagnosis, classification and treatment of diabetes. *British Medical Journal* 2011;342. Available at: http://dx.doi.org/10.1136/bmj.d3319.

19. Stone MA, Camosso-Stenfinovic J, Wilkinson J, de Lusignan S, Hattersley AT and Khunti K. Incorrect and incomplete coding and classification of diabetes: a systematic review. *Diabetic Medicine* 2010;27(5):491–7. Available at: http://dx.doi.org/10.1111/j.1464-5491.2009.02920.x. PMid:20536944.

20. Manuel DG, Rosella LC and Stukel TA. Importance of accurately identifying chronic disease in studies using electronic health records. *British Medical Journal* 2010;341:c4226. Available at: http://dx.doi.org/10.1136/bmj.c4226. PMid:20724404.

21. Gray J, Orr D and Majeed A. Use of read codes in diabetes management in a South London primary care group: implications for establishing disease registers. *British Medical Journal* 2003;326:1130. Available at: http://dx.doi.org/10.1136/bmj.326.7399.1130. PMid:12763987; PMCid:PMC156011.

22. de Lusignan S, Sadek N, Mulnier H, Tahir A, Russell-Jones D and Khunti K. Miscoding, misclassification and misdiagnosis of diabetes in primary care. *Diabetic Medicine* 2011 (Aug) 26. Available at: http://dx.doi.org/10.1111/j.1464-5491.2011.03419.x.

23. de Lusignan S, Khunti K, Belsey J, Hattersley A, van Vlymen J and Gallagher H. A method of identifying and correcting miscoding, misclassification and misdiagnosis in diabetes: a pilot and validation study of routinely collected data. *Diabetic Medicine* 2010 (Feb);27(2):203–9. Available at: http://dx.doi.org/10.1111/j.1464-5491.2009.02917.x. PMid:20546265.

24. Cebul RD, Love TE, Jain AK and Hebert CJ. Electronic health records and quality of diabetes care. *The New England Journal of Medicine* 2011;365(9):825–33. Available at: http://dx.doi.org/10.1056/NEJMsa1102519. PMid:21879900.

25. Kupersmith J, Francis J, Kerr E, Krein S, Pogach L, Kolodner et al. Advancing evidence-based care for diabetes: lessons from the Veterans Health Administration. *Health Affairs* 2007; http://dx.doi.org/10.1377/hlthaff.26.2w156.