Informatics in Primary Care 2004;12:147–56

Problems with primary care data quality: osteoporosis as an exemplar

Simon de Lusignan MSc MRCGP
Senior Lecturer, Primary Care Informatics, Department of Community Health Sciences

Tom Valentin
Medical Student

St George's Hospital Medical School, London, UK

Tom Chan MSc PhD
Research Manager, Surrey and Hampshire Borders NHS Trust, Camberley, Surrey, UK

Nigel Hague BSc MB BCh
Research Fellow

Oliver Wood BSc
Database Officer

Jeremy van Vlymen BEng MSc
Research Assistant

Neil Dhoul BSc
Research Assistant

St George's Hospital Medical School, London, UK

ABSTRACT

Objective To report problems implementing a data quality programme in osteoporosis.

Design Analysis of data extracted using Morbidity Information Query and Export Syntax (MIQUEST) from participating general practices' systems and recommendations of practitioners who attended an action research workshop.

Setting Computerised general practices using different Read code versions to record structured data.

Participants 78 practices predominantly from London and the south east, with representation from north east, north west and south west England.

Main outcome measures Patients at risk can be represented in many ways within structured data. Although fracture data exists, it is unclear which are fragility fractures. T-scores, the gold standard for measuring bone density, cannot be extracted using the UK's standard data extraction tool, MIQUEST; instead manual searches had to be implemented. There is a hundredfold variation in data recording levels between practices. Therapy is more frequently recorded than diagnosis. A multidisciplinary forum of experienced practitioners proposed that a limited list of codes should be used.

Conclusions There is variability in inter-practice data quality. Some clinically important codes are lacking, and there are multiple ways that the same clinical concept can be represented. Different practice computer systems have different versions of Read code, making some data incompatible. Manual searching is still required to find data. Clinicians with an understanding of what data are clinically relevant need to have a stronger voice in the production of codes, and in the creation of recommended lists.

Keywords: computerised medical record, general practice, medical informatics, osteoporosis, primary care, vocabulary, controlled – classification

Introduction

Although United Kingdom (UK) general practice is almost universally computerised, the use of computers to record clinical data is significantly variable.[1,2] The reasons for this are not understood, and as yet there is no widely accepted method to measure the quality of data in general practice computer systems.[3] General practice computer systems have scope for both structured (coded) data and narrative (free text) to be recorded. In the UK, the National Health Service (NHS) requires that general practitioners' (GPs') computer systems should meet certain specifications.[4] One specification is the use of Read codes for the recording of structured data (though this will eventually be replaced by Systematized Nomenclature of Medicine – Clinical Terms [SNOMED-CT]) and that this coded data must be searchable.[5–8] There is no requirement for, or tools provided, to enable the easy analysis of free text. The clinical terminologies, like Read and SNOMED, are getting larger, enabling clinicians to code a wider range of clinical concepts. Read version 2 is the most commonly used classification system in the UK. It comes in two sets, a 4-byte and a 5-byte set. The 4-byte set has about 30 000 terms. However, it has been superseded in most general practice computer systems by the 5-byte code set which offers around 100 000 terms.[9] In 1994, a concept-based coding system was developed (Read 3).[10] This is also known as 'Clinical Terms' and 'CTv3' (Clinical Terms Version 3) and contains over 200 000 clinical concepts (see Table 1). This will not be developed independently in the UK, but instead the CTv3 codes have been merged with the American coding system SNOMED. The new combined version is to be known as SNOMED-CT.[7]

Osteoporosis is a common condition with a high risk of osteoporotic fractures.[11,12] As the population ages the number of fractures is likely to increase, and with it the age-related mortality associated with these fractures.[13,14] Although an important cause of mortality and morbidity, osteoporosis is under-recognised and undertreated, even though DEXA (dual-energy X-ray absorptiometry) scans provide a reliable method of assessing bone density and effective therapy exists.[15–19] The primary care data quality (PCDQ) programme has worked with practices to improve cardiac data quality since 1998.[20] Over the last two years PCDQ has developed an audit-based educational programme in osteoporosis, by applying the techniques developed in other programmes.[21] This programme involved the collection and aggregation of anonymised routinely collected clinical data and its feedback to general practices. The problems with data quality that emerged in the first 78 practices involved in this programme are reported here.

Methods

The first step in the process was to define the key concepts and their relationships – the 'domain ontology' for this condition.[22] A literature review was conducted using PubMed Medline to define the key clinical concepts in osteoporosis. In addition we ran a first action research workshop, the PCDQ Osteoporosis Forum, in order to explore the ways that these

Table 1 Clinical terminologies used in UK general practice computer systems

Name used in text	Read version	Subdivisions of the version	Structure	No. of terms	Commonly used initials
Read Codes	Version 2	4 Byte	Hierarchical	30 000	V2 4-byte
		5 Byte	Hierarchical	100 000	5-byte
Clinical Terms	Version 3	No subdivisions	Concepts and qualifiers	200 000	CTv3
SNOMED-CT (Systematized Nomenclature of Medicine – Clinical Terms) – *Section 2.5*	Amalgamation of Clinical Terms Version 3, and SNOMED RT (Reference Terminology)	No subdivisions	Concepts and qualifiers	>300 000	SNOMED

concepts might be coded (recorded as structured data) by practising clinicians.[23] The participants also asked which patients they would like to have identified as a result of participation in the audit.

The recommendations of the first PCDQ Osteoporosis Forum were that the programme should prioritise the identification of patients according to the following priorities:

- people likely to have fragility fractures, especially those judged at risk of falling; an approach in line with draft national guidance[24,25]
- patients with likely secondary cause of osteoporosis, such as those on steroids, those with diseases likely to cause osteoporosis and those who have had an early menopause[26,27]
- T-scores compatible with osteoporosis or osteopenia.

Fractures of neck of femur, wrist and spine are likely to be fragility fractures, especially if sustained in later adulthood. Our searches looked for these fractures, when they were first coded when the patient was over 40 years old. As osteoporotic fragility fractures are usually the result of a fall from standing height, the management of falls in the elderly is an integral part of reducing the impact of osteoporosis.[28] Hormone replacement therapy (HRT) is prescribed to patients with a premature menopause to prevent osteoporotic fractures, but also to control symptoms of the menopause.

We followed up this workshop with an audit in six practices using the search engines in the clinical systems.[29] The purpose of this was to confirm that the data perceived to be recorded were actually there and to agree an appropriate format for feedback. Finally, we looked at what codes appear at or near the top of the picking lists of the major clinical computer systems when key terms such as 'osteoporosis' are entered (see Figures 1 and 2). Clinicians may be more likely to use codes that appear near the top of the picking list.

The next stage in the development of this programme was the development of Morbidity Information Query and Export Syntax (MIQUEST) queries to extract anonymised data from GP computer systems. MIQUEST was used to extract the data from GP practice databases.[30] MIQUEST allows the same searches to be run on different general practice computer systems. Customised searches/queries were written for the study (NH) using MIQUEST in its 'remote' setting, which allows only anonymised data to be extracted. Only Read-coded data can be extracted using MIQUEST. Free text or narrative data cannot be searched. Therefore information still in paper records, or in text, was not included in the searches. In theory, MIQUEST should run on any GP computer system. However, we found that we needed to customise the queries. We developed separate Read 2 4-byte and 5-byte versions of the query, along with a CTv3 set. In addition, we produced different versions for the Egton Medical Information Systems (EMIS) computer system as it uses codes similar to *British National Formulary* (BNF) chapter headings rather than Read codes for drugs. Pragmatic compromises were made in the development of the queries. Fractures coded for patients once aged over 40 were included in the

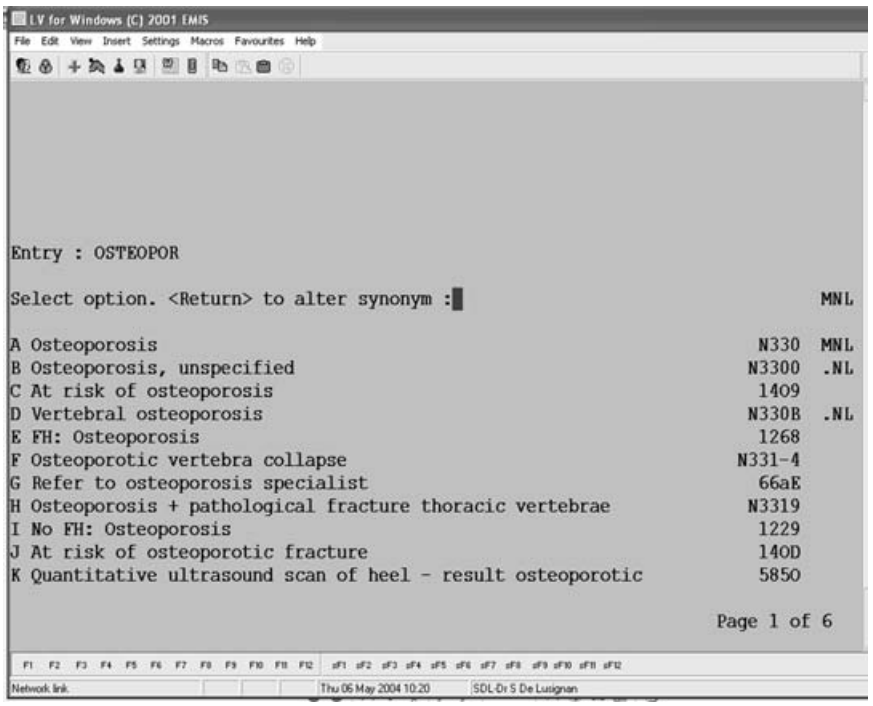

Figure 1 Screenshot of EMIS picking list that appears when 'osteoporosis' is entered as a term

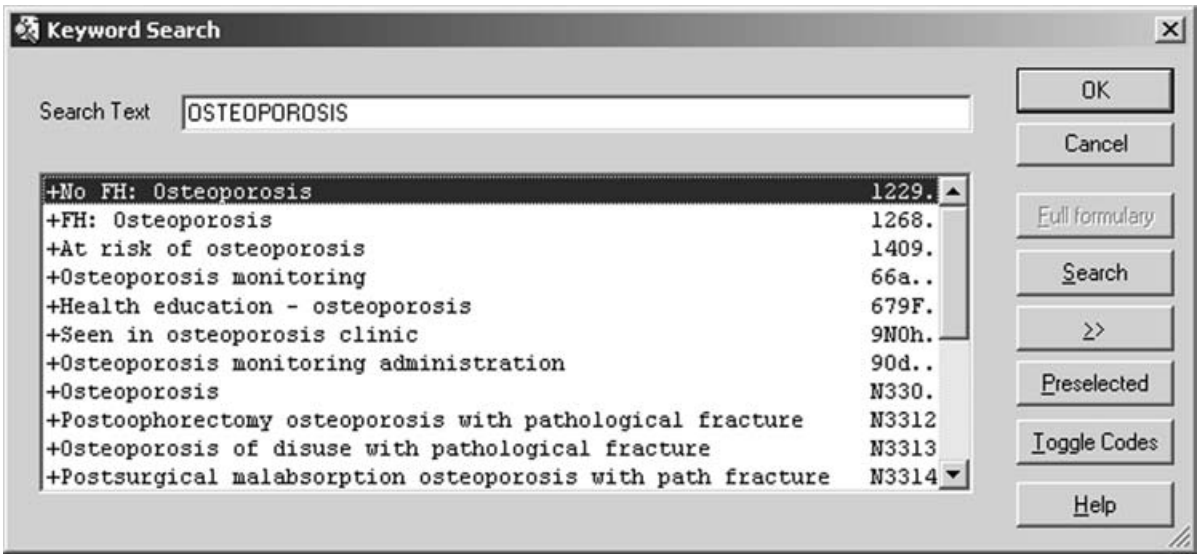

Figure 2 Screenshot of Torex Synergy picking list that appears when 'osteoporosis' is entered as a term

searches as likely fragility fractures even though we are not certain whether they were fractures that occurred earlier in life but have only been coded when the patient was over 40 years. Premature menopause codes were searched for patients for whom the diagnosis was made when <45 years. Chapter headings were used to search for appropriate drugs as it was not considered feasible to run separate searches for each preparation. Read code fh1 was used to identify patients on HRT and on the oestrogen ethinyloestradiol, with additional codes added for other specific treatments, such as code ff8 for tibolone. For EMIS practices the equivalent search was under *BNF* chapter heading 'Oestrogens and HRT' (6.4.1.1.).

Presentations and data summary cards, designed to have most educational impact, were developed based on the pilot data extracted. These were used to feed back data to participating practices and localities. A secure database was set up which contained the practices' names and contact details; it assigned names and numbers to participating practices so that data handled within the office was only identifiable as practice X from Y locality.

Local queries were developed and tested. These create lists of patients who need further structured data recorded. Generally these fall into two groups: patients who are on treatment but have no diagnosis (for example, a patient being treated for osteoporosis who has the details of their diagnosis and scan in a hospital letter), and patients who have a diagnosis suggestive of osteoporosis but no treatment (such as a patient with one or more likely fragility fractures with no treatment prescribed). Brown and Warmington describe the use of Boolean logic to link the presence or absence of one or two data items as a 'data quality probe'.[32]

Volunteer practices took part in this audit. They discussed participation with a member of the research team, received an information pack including details of our information governance policy, and signed a consent form.

The data were exported from the practice systems and imported into a bespoke Microsoft™ Access database (OW, JvV), within which a flat file was created. This was then exported into SPSS (Statistical Package for Social Sciences – Version 12). The process of extracting, aggregating and analysing the data was developed and validated in other disease areas in the context of audit-based education.[21,33,34]

A further action research workshop was held after the first six months of this programme. The anonymised data collected were presented as an action intervention.

Results

Data were extracted from 78 practices in the first six months of the programme. Seventy-one percent (55/78) of the practices had EMIS clinical computer systems, 15% Torex and 8% IPS (In Practice Systems). The 4-byte Read code set was in use in 6.4% of the practices, the rest used 5-byte codes. There were no practices using CTv3 (see Table 2).

Diagnostic recording of osteoporosis, osteopenia and 'at risk of osteoporosis' was known from our previous study to be highly variable between practices. For most practices the number of patients on treatment exceeds the number with a diagnosis. Figure 3 shows the number of patients with osteoporosis compared

Table 2 Cross-tabulation of clinical computer system and Read terminology version

Clinical computer system	Read Version 2		Total
	4-byte	5-byte	
EMIS	1	55	56
Exeter	0	4	4
IPS	0	6	6
Torex	4	8	12
Total	5	73	78

with the number on therapy. Patients on HRT have not been included because this has indications in addition to the treatment of osteoporosis.

It was intended that fragility fractures would be the primary way that we would identify patients with osteoporosis, but there were problems. Crude comparator data for falls, at risk of falls and fracture recording show the lack of pattern of recording between practices. The rates of recording of these are variable, and are set out in detail in Figures 4 to 6. Figure 4 ranks the results by rate of recording of fractures and Figure 5 by rate of recording of falls. A scatter plot (see Figure 6) shows a very weak relationship between practices that record falls and fractures (Pearson Correlation, 2-tailed, weakly significant at the $P<0.05$ level). Although this weak correlation is found, study of Figures 4 and 5 calls into question its validity. The scatter plot also shows that a large number of practices have no falls data, and a smaller number have no fracture data.

Although there are codes for fragility fractures (14G6 history of fragility fracture, and N331 osteoporotic pathological fracture), these codes were hardly used in any of these practices. One practice represented at the Forum used the Austin Moore hip replacement code (7K23 or 7K24) to indicate whether a hip fracture was pathological or not. However, they often guessed between these codes as the difference between them is whether or not the orthopaedic surgeon used cement in the operation. This information is not usually available in the discharge summary sent to practitioners.

Recording of secondary causes of osteoporosis was much more consistent, especially steroid prescription data (data not shown). However, data about premature menopause and HRT were much more confused. Premature menopause can be represented by a large number of codes. The commonest codes identified were: excision of uterus (7E04), menopause monitoring (66U) and absence of menstruation (K590). However, some practices used other codes, for example: menopause (1512), artificial menopause (K5A4) and premature menopause (C1613). The MIQUEST processor did not read the *BNF* chapter heading 6.4.1.1. correctly – it ignored the last digit and searched instead on chapter 6.4.1. 'Sex hormone', which included in addition 6.4.1.2. 'Progestogens', thereby inflating the numbers of patients on therapy. HRT prescription recording varies greatly between practices (see Figure 7), and box plots show that the 'HRT' queried captured more data from EMIS practices (see Figure 8). If outliers with apparently very high levels of prescribing

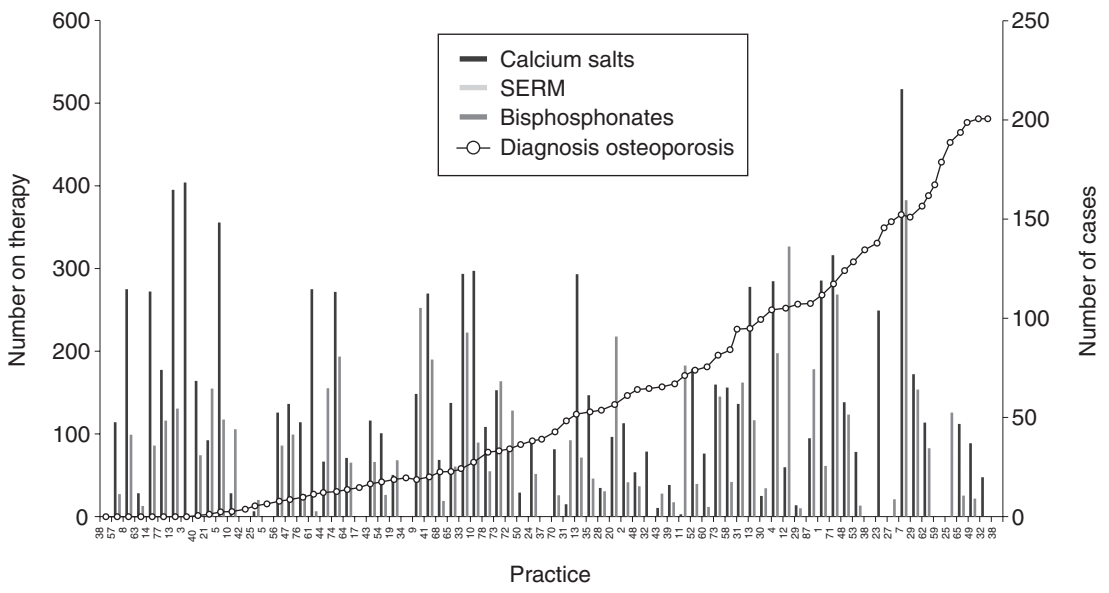

Figure 3 In nearly all practices the number of patients receiving treatment exceeds the number with a diagnosis

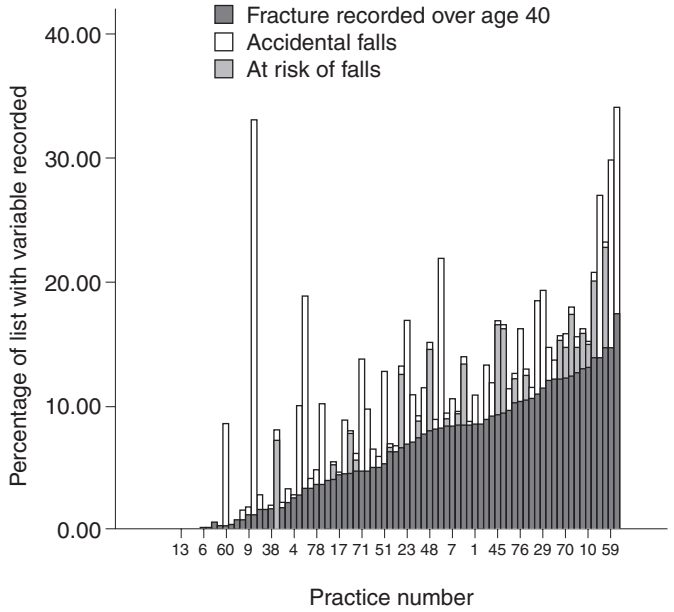

Figure 4 The 78 practices in the audit ranked by rate of fracture recording. Recordings of falls and 'at risk of falls' are superimposed, illustrating the variation between practices

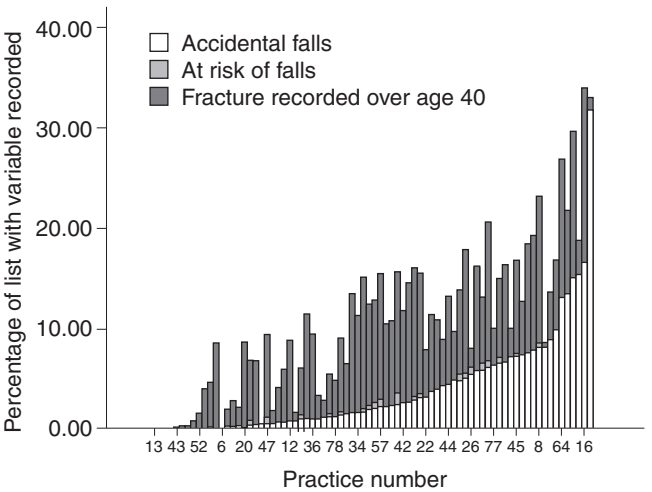

Figure 5 The 78 practices in the audit ranked by rate of recording of accidental falls. Recordings of fractures and 'at risk of falls' are superimposed, illustrating the variation between practices

are excluded, the mean rate of prescription in EMIS practices was 13.2% (standard deviation 5.5%) and for non-EMIS practices 9.3% (SD 4.8%). This difference is significant at the $P<0.05$ level (t-test). The box of the box-whisker indicates the first and third quartiles and is divided by the median. The whiskers extend up to 1.5 times the interquartile range from the box to include the range.

Numeric T-scores could not be extracted from any of the practice computer systems. Whilst there are codes for T-scores, none of the practices had any numeric data associated with such codes. This was because the numeric field associated with a given Read code sometimes needs to be a negative number. The gold standard for making a diagnosis of osteoporosis is a T-score of less than –2.5. T-scores are negative numbers, and the clinical systems would not allow the entry of the minus sign into a numeric field.

The PCDQ Osteoporosis Forum commented on the need for a short recommended list of codes: standardisation would make it easier to find patients when searches were run. There needs to be a way of entering numeric T-scores which accommodates the need for many of these scores to be negative numbers. The codes need to be revised so that how the T-score is measured, and their numerical values, are taken into account.

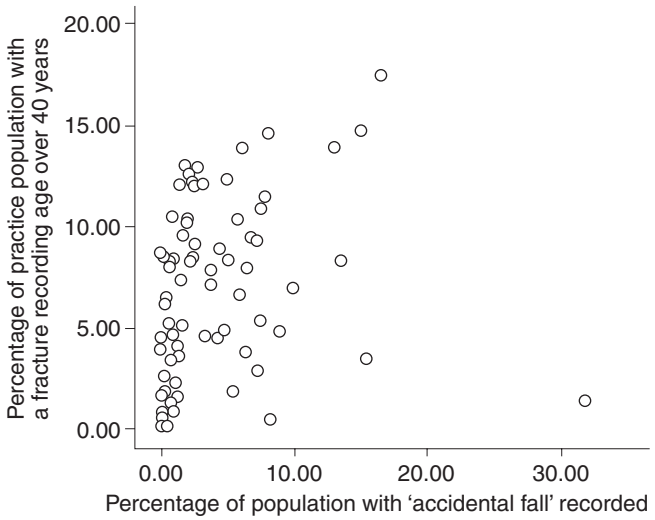

Figure 6 Scatter plot comparing rates of recording of fractures and falls ($P<0.05$)

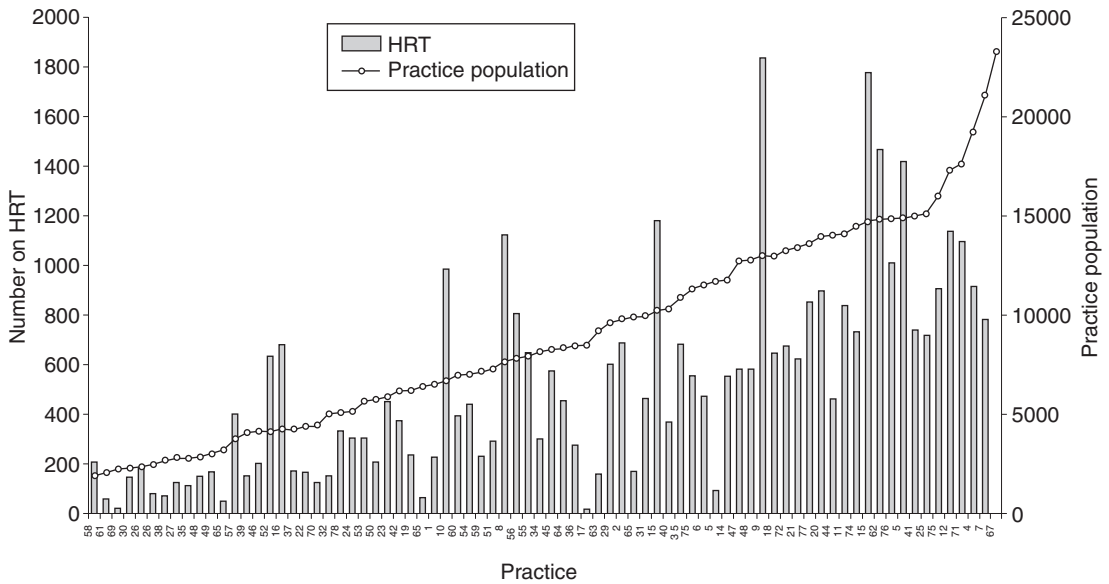

Figure 7 Inter-practice variation in 'HRT' prescribing

Discussion

These data show that there is much to be done to improve data quality in osteoporosis. The problems identified provide new insights and generalisable lessons about structured data and data quality. It is impossible to record numeric T-scores, the gold-standard diagnostic test, as structured clinical data because computer systems will not accept negative numbers into their numeric fields. It is also not possible to characterise fractures as osteoporotic (or as fragility fractures) in the way that they can readily be coded as open or closed. Codes that provide qualitative information, such as T-score osteoporotic, do exist. However, in other clinical areas the use of this type of code – for example, cholesterol raised – has proved less useful than coding the numerical value as thresholds for diagnosis and treatment change over time. There are also no codes to record the type of machine that was used to perform the scan, something that influences the validity of the result.[35,36] Having a large number of ways in which a clinical concept can be recorded gives plenty of licence to the clinician, but makes finding these patients extremely difficult. Having a number of variants of a clinical terminology makes data extraction more time-consuming and expensive, and aggregation of data imprecise. Selecting codes from non-standard picking lists further exacerbates this situation as clinicians under time pressure in the consultation may tend to select a code on the first page and near the top of the

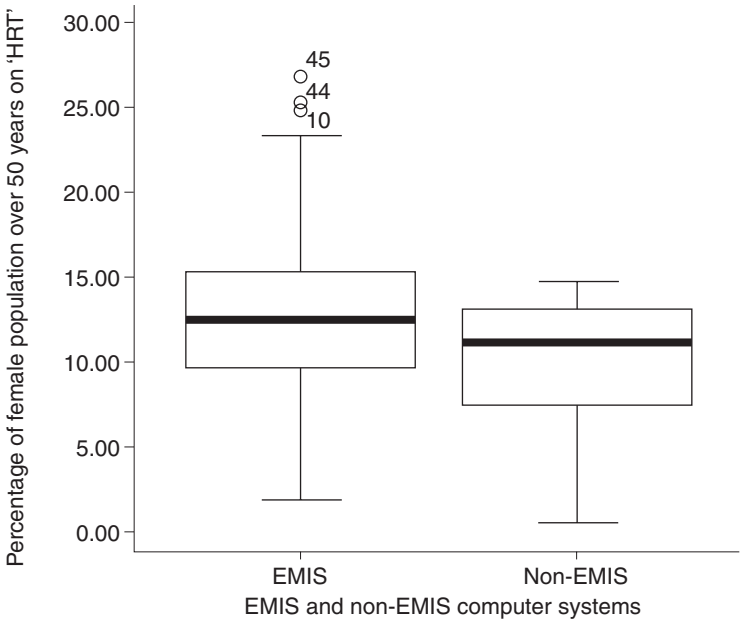

Figure 8 Box-whisker plots showing the variation in rates of prescribing of 'HRT' between EMIS and non-EMIS practices

list. Technical constraints restrict the effectiveness of data extraction tools and sometimes searches fail. Primary care professionals with an interest in this disease area believe that there should be a limited list of codes, which are designed to enable patients with clinical need to be readily identified from clinical records.

These findings suggest that as coding and classification systems get larger there are more and more ways in which a concept can be coded; this risks making it harder to find patients who may benefit from interventions. The lack of retrievable numeric T-scores means that manual searching of data remains the only way to find this information. The migration to SNOMED-CT provides an opportunity to standardise the use of codes across clinical computer systems, removing the difficulties associated with the use of different variants of the same coding system.[7] This could also provide the opportunity to standardise the picking lists of codes. Finally, clinicians need protected time to share and explore reasons for inter-practice variation.

It is possible that we did not find all the ways that relevant data were coded. We would not have looked for Austin Moore hip replacement codes had a member of the Forum not explained that this is the code his practice uses to represent a fragility fracture of the hip. Data, including T-scores, might have been recorded in the free-text record, including scanned-in hospital letters and reports, and therefore have been invisible to our searches. Searches and data extractions are not always successful, though we have excluded failed searches from the data – where zero data appears it is where no data was recorded.

Rector has described the problems with reusing data recorded for one purpose for another.[37] In this case the reuse of routinely collected data, in order to improve the quality of care, was challenging because different practices represented the same clinical concepts differently within the coding system. Consistent data recording within practices, alongside marked inter-practice variation, has been observed for over ten years.[38] As early as 1992, work was published illustrating variability in coding habits. Subsequently, Fleming *et al*, through the network of the Royal College of General Practitioners' Birmingham Research Unit (RCGP–BRU), developed a hierarchy for coding problem titles that is designed to encourage greater coding conformity.[39,40] Variation has also been reported for osteoporosis data recorded between 1994 and 1998.[41] This study demonstrates that despite enormous steps forward in other areas of data quality, data quality in osteoporosis remains poor.[21,34,42] The lack of any link between diagnosis and therapy in most computerised records makes it impossible to understand why a particular prescription has been issued. Carey *et al* have described the advantages of problem-orientated medical records (POMR).[43] Had the study practices had POMR it would have helped. For example, we could have differentiated between HRT issued for postmenopausal symptoms and that prescribed to protect against osteoporotic fractures. Gray *et al* have looked in heart disease at the positive predictive value and sensitivity of diagnostic codes and therapy for actually having the condition.[44] This is similar to our usual approach to implementing quality improvement, that is, to identify a subset of patients with the

diagnosis or on treatment and then to examine key risk factors and associated co-morbidities.[45] However, the low rate of diagnostic recording, and the variation between practices in their choice of therapy, precluded us from doing this in this case.

Further research is needed to explore the reasons for inter-practice variation in coding and whether there is an optimum size for a clinical terminology.

Conclusion

Clinical terminologies are being developed that provide many ways in which the same clinical concept can be represented, making comparable information hard to find and aggregate. Despite the growth in the size of terminologies, some clinically important data cannot be coded. Coding and therapeutic choices appear to be consistent within practices. The principal level of variation is between practices. The existence of several different versions of a clinical coding system, and presenting clinicians with different picking lists, are further barriers to raising data quality standards. The data quality issues that have arisen in the context of this quality improvement programme have relevance beyond osteoporosis.

ACKNOWLEDGEMENTS

The PCDQ osteoporosis programme was supported by MSD, through an unconditional educational grant, January 2002 to March 2004. We thank the volunteer practices who took part in the audit and in the Osteoporosis Forum. Tom Valentin used his student special study module to work on this paper.

REFERENCES

1 Department of Health. *Primary Care Computing*. www.dh.gov.uk/PolicyAndGuidance/OrganisationPolicy/PrimaryCare/PrimaryCareComputing/fs/en
2 Munday E. Data quality issues plague primary care. *British Journal of Healthcare Computing and Information Management* 2003;20(3):22–4.
3 Thiru K, Hassey A and Sullivan F. Systematic review of scope and quality of electronic patient records in primary care. *British Medical Journal* 2003;326:1070–2.
4 NHS Information Authority. NAPPS (National Accreditation and Procurement Process Service). www.nhsia.nhs.uk/napps/pages/default.asp
5 NHS Information Authority. *Clinical Terminology Service*. www.nhsia.nhs.uk/terms/pages/readcodes_intro.asp
6 Severs MP. The Clinical Terms Project. *Bulletin of the Royal College of Physicians of London* 1993;17(2):9–10.
7 NHS Information Authority. *SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms)*. www.nhsia.nhs.uk/snomed/
8 Stearns MQ, Price C, Spackman KA and Wang AY. SNOMED clinical terms: overview of the development process and project status. In: Bakken S (ed). *Proceedings of the 2001 AMIA Fall Symposium*, pp. 662–6. Philadelphia, PA: Hanley and Belfus, 2001.
9 Robinson D, Wanger K and Price C. The clinical Philadelphia terms and ICPC: identifying equivalence and enabling compatibility for 'non-medical' terms. British Computer Society, *Proceedings of the Primary Care Specialist Group Annual Conference 1998*. phcsg.ncl.ac.uk/conferences/cambridge1998/robinson.htm
10 O'Neil M, Payne C and Read JD. Read codes version 3 – a user-led terminology. *Methods of Information in Medicine* 1995;34:187–92.
11 Cooper C. Epidemiology of osteoporosis. *Osteoporosis International* 1999;Suppl.2:S2–S8.
12 Porthouse J, Birks YF, Torgerson DJ, Cockayne S, Puffer S, Watt I. Risk factors for fracture in a UK population: a prospective cohort study. *QJM* 2004;97(9):569–74.
13 Cooper C, Campion G and Melton IJ. Hip fractures in the elderly: a worldwide problem. *Osteoporosis International* 1992;2:285–9.
14 Roberts SE and Goldacre MJ. Time trends and demography of mortality after fractured neck of femur in an English population, 1968–98: database study. *British Medical Journal* 2003;327:771–5.
15 Saadi H, Litaker D, Mills W, Kippes C, Richmond B and Licata A. Practice variation in the diagnosis and treatment of osteoporosis: a case for more effective physician education in primary care. *Journal of Women's Health Gender-Based Medicine* 1999;8(6):767–71.
16 Chestnut CH. Osteoporosis, an underdiagnosed disease. *Journal of the American Medical Association* 2001;286:2865–6.
17 Siris ES, Miller PD, Barrett-Connor E *et al*. Identification and fracture outcomes of undiagnosed low bone mineral density in postmenopausal women: results from the National Osteoporosis Risk Assessment. *Journal of the American Medical Association* 2001;286:2815–22.
18 Miller PD, Siris ES, Barrett-Connor E *et al*. Prediction of fracture risk in postmenopausal white women with peripheral bone densitometry: evidence from the National Osteoporosis Risk Assessment. *Journal of Bone Mineral Research* 2002;17:2222–30.
19 Delmas PD. Osteoporosis IV: treatment of postmenopausal osteoporosis. *Lancet* 2002;359:2018–26.
20 de Lusignan S and Hague NJ. The PCDQ (Primary Care Data Quality) Programme. *Bandolier/Impact* 2001; January. www.jr2.ox.ac.uk/Bandolier/booth/mgmt/PCDQ.html
21 de Lusignan S, Hague N, Brown A and Majeed A. An educational intervention to improve data quality in CHD. *Journal of Public Health* 2004;26(1):34–7.
22 Musen MA. Medical informatics: searching for underlying components. *Methods of Information in Medicine* 2002;41:12–19.
23 Hart E and Bond M. *Action research for health and social care: a guide to practice*. Buckingham: Open University Press 1995.
24 National Institute for Clinical Effectiveness (NICE). *Technologies for the Prevention and Treatment of Osteoporosis and Prevention of Osteoporotic Fractures in*

Postmenopausal Women [appraisal consultation document]. www.nice.org.uk/article.asp?a=97559
25 National Institute for Clinical Effectiveness (NICE). *Technology Appraisal: osteoporosis prevention and treatment*. www.nice.org.uk/cat.asp?c=20318
26 Royal College of Physicians. *Osteoporosis: clinical guidelines for prevention and treatment*. London: Royal College of Physicians, 1999.
27 Royal College of Physicians. *Osteoporosis: clinical guidelines for prevention and treatment. Update on pharmacological interventions and an algorithm for management*. London: Royal College of Physicians, 2000. www.rcplondon.ac.uk/pubs/wp_osteo_update.htm
28 Woolf AD and Akesson K. Preventing fractures in elderly people. *British Medical Journal* 2003;327:89–95.
29 de Lusignan S, Chan T, Wells S *et al*. Can the patients with osteoporosis, who should benefit from the implementation of the National Service Framework for Older People, be identified from general practice computer records? A pilot study that illustrates the variability of computerised medical records and problems with searching computerised medical records. *Public Health* 2003; 117(6):438–45.
30 St. George's Hospital Medical School. *Primary Care Data Quality Project. What is MIQUEST*? www.pcdq.org/MIQUEST.htm
31 Egton Medical Information Systems (EMIS): www.emis-online.com/
32 Brown PJ and Warmington V. Info-tsunami: surviving the storm with data quality probes. *Informatics in Primary Care* 2003;11(4):229–33.
33 de Lusignan S, Dzregah B, Hague N and Chan T. Cholesterol management in patients with ischaemic heart disease: an audit-based appraisal of progress towards clinical targets in primary care. *British Journal of Cardiology* 2003;10:223–8.
34 de Lusignan S, Belsey J, Hague N and Dzregah B. End digit preference in blood pressure recordings of patients with ischaemic heart disease in primary care. *Journal of Human Hypertension* 2004;18:261–5.
35 Pocock NA, Culton NL, Gilbert GR *et al*. Potential roles for quantitative ultrasound in the management of osteoporosis. *Medical Journal of Australia* 2000;173(7): 355–8.
36 Moyad MA. Osteoporosis: a rapid review of risk factors and screening methods. *Urologic Oncology: Seminars and Original Investigations* 2003;21(5):375–9.
37 Rector A. Terminology, codes and classifications in perspective: the challenge of reuse. *British Journal of Healthcare Computing and Information Management* 2000;7(3):20–3.
38 Crombie DL, Cross KW and Fleming DM. The problem of diagnostic variability in general practice. *Journal of Epidemiology and Community Health* 1992;46(4):447–54.
39 Fleming DM. The measurement of morbidity in general practice. *Journal of Epidemiology and Community Health* 1991;45(3):180–3.
40 Fleming DM. Weekly Returns Service of the Royal College of General Practitioners. *Communicable disease and public health/PHLS* 1999;2(2):96–100.
41 Bailey K, Ellis C and Majeed A. Trends in the treatment of osteoporosis and types of drug treatments prescribed in general practice in England and Wales, 1994 to 1998. *Health Statistics Quarterly* 2002;16:5–13.
42 Horsfield P. Trends in data recording by general practice teams: an analysis of data extracted from clinical computer systems by the PRIMIS project. *Informatics in Primary Care* 2002;10:227–34.
43 Carey IM, Cook DG, De Wilde S *et al*. Implications of the problem-orientated medical record (POMR) for research using electronic GP databases: a comparison of the Doctors' Independent Network Database (DIN) and the General Practice Research Database (GPRD). *BMC Family Practice* 2003; September 30.
44 Gray J, Majeed A, Kerry S and Rowlands G. Identifying patients with ischaemic heart disease in general practice: cross-sectional study of paper and computerised medical records. *British Medical Journal* 2000;321:548–50.
45 Thiru K, de Lusignan S, Sullivan F, Brew S and Cooper A. Three steps to data quality. *Informatics in Primary Care* 2003;11(2):95–102.

CONFLICTS OF INTEREST

None.

ADDRESS FOR CORRESPONDENCE

Simon de Lusignan
Senior Lecturer, Primary Care Informatics
Department of Community Health Sciences
St George's Hospital Medical School
London SW17 0RE
UK
Tel: +44 (0)20 8725 5661
Fax: +44 (0)20 8767 7697
Email: slusigna@sghms.ac.uk

Accepted July 2004