

Quality of morbidity coding in general practice computerized medical records: a systematic review

Kelvin Jordan, Mark Porcheret and Peter Croft

Jordan K, Porcheret M and Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. *Family Practice* 2004; **21**: 396–412.

Background. Increased use of computers and morbidity coding in primary care delivery and research brings a need for evidence of the quality of general practice medical records.

Objective. Our aim was to assess the quality, in terms of completeness and correctness, of morbidity coding in computerized general practice records through a systematic review.

Methods. Published studies were identified by searches of electronic databases and citations of collected papers. Assessment of each article was made by two independent observers and discrepancies resolved by consensus. Studies were reviewed qualitatively due to their heterogeneity.

Results. Twenty-four studies met the inclusion criteria for the review. There was variation in the methodology and quality of studies, and problems in generalizability. Studies have attempted to assess the completeness and correctness of morbidity registers by reference to a gold standard such as paper notes, prescribing information or diagnostic tests and procedures, each of which has problems. A consistent finding was that quality of recording varied between morbidities. One reason for this may be in distinctiveness of diagnosis (e.g. coding of diabetes tended to be of higher quality than coding of asthma).

Conclusions. This review highlights the problems faced in assessing the completeness and correctness of computerized general practice medical records. However, it also suggests that a high quality of coding can be achieved. The focus should now be on methods to encourage and help practices improve the quality of their coding.

Keywords. Medical records, primary care, systematic review.

Introduction

Computers in primary care are used for clinical management, administration, research and planning. A 1996 survey of general practices in England showed that 96% were computerized and 81% of these used computers for entering clinical data during consultations.¹ A systematic review of published studies on primary care computing concluded that most GPs accepted computers in their working lives, thought they were more accurate than paper notes, gave them better access to records and improved patient care.²

Morbidities can be entered onto the computer using various coding classifications. One of the more common in the UK is the Read Code classification, a hierarchy of morbidity, symptom and process codes which become more specific further down the hierarchy.³ However, the

quality of consultation recording and morbidity coding needs to be established. One review of the quality of computer-held patient records mainly assessed studies of hospital records.⁴ A recent review of how the quality of primary care computerized records (including morbidity, prescription, referral, lifestyle and socio-economic data) has been assessed concluded that there was a lack of standardized methods and that recorded prescription data appeared to have the highest overall quality.⁵

Our objective was to conduct a review focused on morbidity data in order to assess in detail the quality of recording of morbidity codes and morbidity registers held in computerized primary care records in the UK and to evaluate the different approaches used to assess this quality.

Methods

We define a morbidity register as a list of people with a certain morbidity code or disease. This does not have to be an actual computerized list, as it may also be

Received 19 March 2003; Revised 27 August 2003; Accepted 10 March 2004.

Primary Care Sciences Research Centre, Keele University, Keele, Staffs ST5 5BG, UK; E-mail: k.p.jordan@keele.ac.uk

established by searching the electronic records for all patients with a particular morbidity code or codes.

Defining quality

The quality of coding is often described by reference to its 'accuracy', but defining this term is difficult. The accuracy of morbidity coding is subjective, and can often only be established by reference to a gold standard.

The review by Hogan and Wagner assessed accuracy in terms of the completeness and correctness of records.⁴ Completeness was defined as equivalent to sensitivity: the proportion of observations 'about the world' that were actually recorded. For example, with respect to a morbidity register, are all the subjects with the morbidity included on the register? Correctness was defined as equivalent to the positive predictive value (PPV): the proportion of observations that reflect the 'true state of the world'. For example, do all subjects on a register actually have that disease? Both are necessary for a database to be accurate. A high level of correctness may be achieved at the expense of failing to record all information, i.e. poor completeness.⁴ Similarly, a high level of completeness may be obtained at the cost of poor correctness.

In this review, we expand these definitions and examine four criteria. (i) The completeness of consultation recording—for each contact a patient has with the GP, is there a morbidity code recorded on the computer? This is an important element of completeness because, if no code is allocated or if the contact goes unrecorded, then completeness of the database is compromised. Further, is each different clinical morbidity consulted about within one contact coded? (ii) The correctness of consultation recording—are the codes given during this contact appropriate? (iii) The completeness of a morbidity register—is everyone included on the register that should be? (iv) The correctness of a morbidity register—should everyone on a register be on that register?

Criterion (i) can normally be evaluated by the percentage of consultations recorded on computer and the percentage of these with a morbidity code. Criteria (ii)–(iv) can be split into external and internal validation.

- (a) Does the patient actually have the morbidity indicated by the code given? This is examined by reference to external sources such as experts in the field of that morbidity examining the patient, and relates to the diagnostic abilities of the GP. This will be referred to as external validity.
- (b) Is the code given accurate based on the evidence that is available within the primary care practice or clinic? This could be evidence obtained during a consultation, or information from external sources supplied to the practice, i.e. is the code the one which would be expected given information such as

the history of the patient, hospital letters and test results or prescriptions issued by the GP. This could also relate to whether the code given is correct based on what morbidity the GP thinks the patient has. This will be referred to as internal validity.

In this review, we examine the four criteria above in relation to internal validity. The objective of the review is not to assess the diagnostic ability of GPs, but to assess whether diagnoses in primary care electronic records are a true and complete reflection of the diagnoses given by the practitioner and the information available to the practitioner.

Criteria for inclusion of studies in the review

Inclusion criteria.

- (i) Studies should be based in primary care.
- (ii) Studies should assess computerized records or a computerized morbidity register (studies which did not explicitly state that the records were computerized were excluded).
- (iii) Studies should be based in the UK. This allows us to assess the quality of recording in the UK without disentangling the effects of different primary care systems.
- (iv) Studies should have a stated objective of assessing the quality of these records based on criteria (i)–(iv) above.

Exclusion criteria.

- (i) Studies which solely compare disease prevalence rates with external (to the practice) rates.
- (ii) Studies which attempt to validate general practice diagnosis by reference to external sources such as experts in the field of that morbidity reviewing the patient (i.e. external validity).
- (iii) Studies which use fictional or simulated patients. These represent another method to investigate external validity.
- (iv) Studies comparing patient self-reported consultation with medical records. Here, it is uncertain whether the medical records are validating the self-report, or vice versa.

Databases searched were MEDLINE, Science Citation Index, Social Science Citation Index, Cumulative Index to Nursing and Allied Health Literature (CINAHL), English National Health Care database, the Cochrane Library and the National Research Register. Citations of collected articles were also searched. English language articles up to September 2002 were collected. Keywords used were at three levels, with articles examined for at least one word in its title, abstract or keywords from each level. Level one keywords were: 'primary care', 'general pract*', 'family pract*', level two were 'morbid*',

'computer*', 'record*', 'electronic', 'register', 'consult*', 'contact*'; and level three were 'agree*', 'valid*', 'accura*', 'complete*', 'correct*', 'reliab*'.

Following the search strategy, studies which obviously failed to meet the inclusion/exclusion criteria (based on their abstract) were discarded and the full papers of the remainder read. Studies which fitted the inclusion and exclusion criteria were included in the review.

Two independent observers assessed each included article using a data extraction sheet. Discrepancies between the observers were resolved through consensus. One observer (KJ) is a biostatistician, the other (MP) is a practising GP. Due to the heterogeneity of the studies, no formal meta-analysis or pooling of the data was possible. Studies were reviewed qualitatively.

Results

Following the search strategy, 344 potential studies were identified. After assessment of their abstracts, 89 papers were read and 24 included in the review. Reasons for exclusion are given in Figure 1.

A summary of the methods of the studies in the review are given in Table 1 with their results in Table 2.

Methodologies of the studies

Three studies were single practice evaluations⁶⁻⁸ and eight studies evaluated the VAMP database system used for the General Practice Research Database (GPRD).⁹⁻¹⁶ Two other studies assessed practices contributing towards a multipractice database; the MediPlus database¹⁷ and the GPASS system in Scotland.¹⁸ The number of practices for the remaining 11 studies ranged from two to 25 (median 7). After VAMP, the most common computer system used was EMIS (seven studies).^{6,8,19-23} No more than two studies used any other system. Only six studies were published before 1995.^{7,9-11,24,25}

All the VAMP studies used OXMIS for morbidity coding. Ten studies had practices stated to be using Read Codes,^{6,8,17-19,21-23,26,27} whilst two studies used ICD9.^{11,25}

The paper information in the practice was the most commonly used gold standard (nine studies).^{6,8,11,18-22,27} This could include, for example, hospital letters as well as notes about the consultation kept in these manual notes. Prescription information (paper or electronic) on drugs relating to a specific disease was used as the sole gold standard for at least one quality assessment by eight studies.^{6,13,15,21-24,28} Combinations of disease-specific information stored electronically,^{6,12} information including drug records and hospital letters in paper format¹⁸ or information stored either electronically or in paper notes^{21,26,29} were also used. For example, Hassey *et al.* used drugs and, where appropriate, other diagnostic tests stored electronically

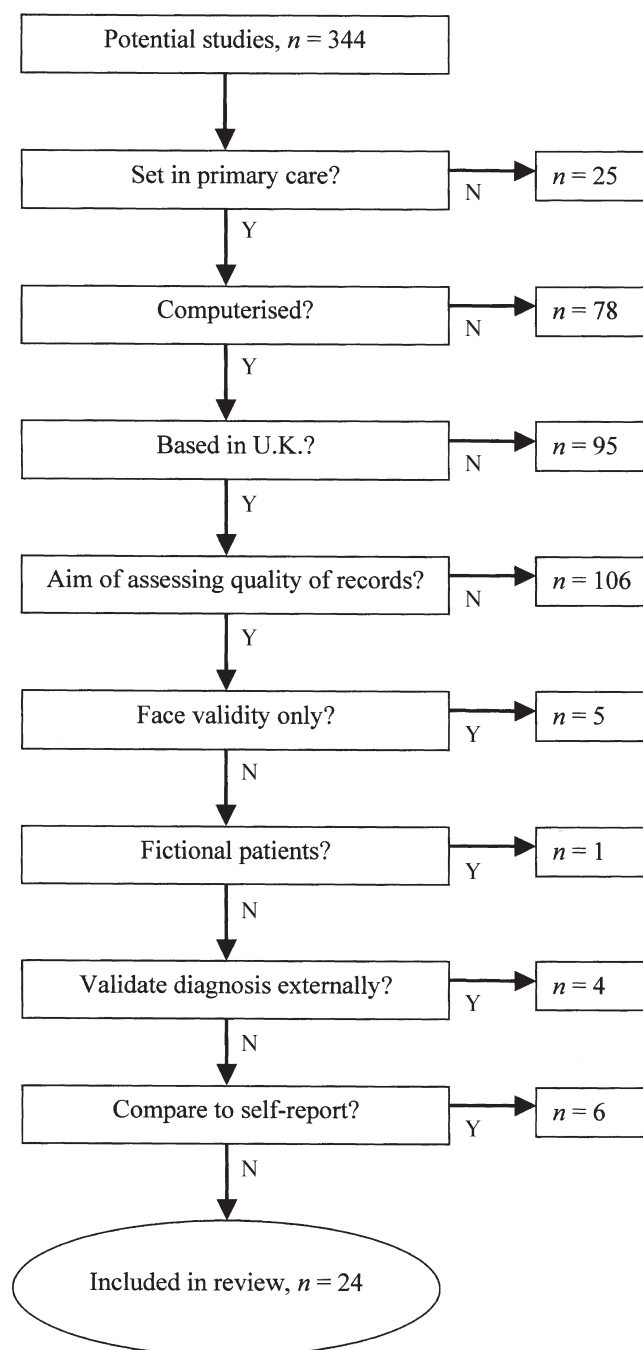


FIGURE 1 Inclusion and exclusion criteria for the review. A study excluded for a reason near the top of the flowchart may also have failed to meet later criteria

to validate 15 diagnoses.⁶ Pringle *et al.* used a combination of paper notes, computerized prescription data and related diagnoses stored on computer.²¹ They also videotaped consultations to assess correctness of consultation coding compared with independent coding. Several GPRD studies sent questionnaires to the patients' GPs asking for information about symptoms or reason for diagnosis^{12,15,16} or to ask for hospital correspondence.^{9,10,14} One study used hospital discharge information maintained by the Health Authority.²⁵

TABLE 1 Methodology of the studies in the review

Study	Computer system/ software	Coding system	Quality of consultation reporting	Completeness of morbidity register	Correctness of morbidity register	What assessed	'Gold standard'/ assessment criteria	No. of records/ patients/GPs/ practices assessed
Mant and Tulloch (1987) ²⁵ Oxford	Not stated	ICD9		✓		Disease registers for cancer, diabetes, MI, epilepsy, hypertension, thyroid disease	Hospital Activity Analysis data (patients discharged from hospital with disease) for 1982–1984	4 practices—2 entered data weekly onto project's computer, 2 entered data interactively at practice
Coulter <i>et al.</i> (1989) ²⁴ Oxford	Mainly (6/7 practices) VAMP	Not stated		✓		Chronic disease registers for epilepsy, diabetes, thyroid disorders and asthma.	Repeat prescriptions	7 practices with chronic disease registers or patient summaries and used computers for repeat prescriptions
Gilliland <i>et al.</i> (1992) ⁷ Northern Ireland	Not stated	Not stated	✓			Recording of problems and diagnoses from consultation	Monthly percentage of consultations with a recorded problem/diagnosis	1 practice: GPs record diagnosis on encounter form and clerks enter it onto computer. 2 groups: study group (5 GPs) who received 3 interventions 4 months apart: (i) 1 to 1 request to record diagnosis; (ii) 1 to 1 request with feedback of own 3 monthly results; (iii) group feedback of results. Control group (4 GPs) had no intervention
Pringle <i>et al.</i> (1995) ²¹ Trent Region	EMIS	Read Codes	✓	✓	✓	(i) Diabetes, glaucoma, asthma and CHD computer morbidity registers. (ii) and (iii) Coding of consecutive consultations	(i) Paper notes, repeat prescribing and related diagnoses for diabetes and glaucoma; for asthma and CHD % on register who received asthma and nitrate medication, respectively, in previous 2 years. (ii) Paper notes. (iii) Videotape of consultations reviewed by 1 of 2 observers (reliability tested in pilot study).	(i) 4 practices committed to recording. (ii) 50 consecutive consultations for each GP in each practice (iii) 50 consecutive consultations of each of 4 GPs in 1 practice

TABLE 1 *Continued*

Study	Computer system/ software	Coding system	Quality of consultation reporting	Completeness of morbidity register	Correctness of morbidity register	What assessed	'Gold standard'/ assessment criteria	No. of records/ patients/GPs/ practices assessed
Scobie <i>et al.</i> (1995) ²² London	3 Meditel, 3 EMIS	Read Codes	✓	✓		(i) Whether consultation entered on computer. (ii) Whether consultation coded. (iii) Diabetes, asthma, angina and epilepsy computer registers.	(i) Paper notes reviewed by single researcher after . pilot (1 year); (ii) % of reasons for consulting given Read Code (1 year); (iii) Prescribed medication	(i) 915 patients from 5 practices; (ii) 405 patients from 3 practices of whom 284 consulted during year; (iii) patients from 6 practices
Neal <i>et al.</i> (1996) ²⁰ Yorkshire	EMIS	N/A	✓			Extent of recording of consultations	Paper notes	4 practices with 'required data' out of initial 38, 'random sample' of patients; no. of consultations per practice ranged from 1065 to 2920
Pearson <i>et al.</i> (1996) ²⁷ Somerset	Not stated	Read Codes	✓			(i) Level of coding of consultations (ii) 'Appropriateness' of code	Paper notes examined by project co-ordinator.	(i) 11 practices over 1 year. (ii) 25 patients every 3 months for 1 year for each practice
Whitelaw <i>et al.</i> (1996) ¹⁸ Scotland	GPASS	Read Codes		✓	✓	19 diagnoses	Gold standards were combinations of paper records, clinical summary sheet, hospital letter, drug record (which were all assessed by one trained field worker) and patient questionnaire.	250 patients (aged 45–64) per practice selected from 41 practices with >50% of patients with clinical Read Code, out of 410 practices in GPASS project, leading to 5567 (54%) patients with complete information
Bottomley (1997) ²⁸ Wakefield	Not stated	Not stated		✓	✓	Prevalence of angina based on CHD register	Prevalence of angina based on nitrate prescription (patients with >1 nitrate prescription in 1992)	38/48 practices contacted responded, 21 of these able to provide information, 9 had CHD register with usable information
Hooker <i>et al.</i> (1999) ⁸ London	EMIS	Read Codes		✓	✓	Coding of hypertension	Paper notes reviewed by 1 researcher	Random sample of 100 records for elderly patients in 1 practice
Thiru <i>et al.</i> (1999) ²³ Surrey	EMIS/ AAH Meditel	Read Codes		✓	✓	Consultation prevalence of asthma, diabetes, glaucoma and CHD diagnoses	Prevalences based on medication and, for glaucoma, trabeculectomy	2 practices (1 Meditel, 1 EMIS), all patients registered from 1/1/1994 to 1999

Gray <i>et al.</i> (2000) ¹⁹ Battersea	EMIS	Read Codes		✓	✓	Different computer search strategies using G3 Read Code with drug codes (nitrate, aspirin, atenolol, digoxin, statin) to obtain cases of IHD	Compared with paper case notes. Definite case if report of diagnostic ECG, raised cardiac enzyme activity confirming MI or positive coronary angiogram, exercise test or thallium scan. Probable case if written record strongly suggested IHD and angina drugs prescribed	11 practices, 1680 patient records (1 in 40 random sample of patients aged 45 and over) examined leading to 80 definite and 13 probable cases
Moher <i>et al.</i> (2000) ²⁹ Warwickshire	Not stated (1 practice not computerized)	Not stated		✓	✓	Patients on CHD morbidity register,	Information from hand search of paper notes and computer by 6 external auditors for MI, stable angina, percutaneous transluminal coronary angioplasty or coronary artery bypass grafting. Defined by ECG, cardiac enzyme changes, relevant history or hospital discharge letter	18 practices with CHD morbidity register patients aged 55–75 leading to 1979 cases
Hassey <i>et al.</i> (2001) ⁶ North Yorkshire	EMIS	Read Codes	✓	✓	✓	(i) Extent of consultation recording (ii) Validation of 15 diagnostic Read Codes.	(i) List of appointments for 1 week. (ii) Other relevant electronically recorded criteria for that diagnosis (drugs, other treatments, tests)	1 practice
Connolly <i>et al.</i> (2002) ²⁶ Northern Ireland	Various	Read Codes			✓	Cases of CHD found by computer search with any of: (i) nitrate prescription in previous 12 months; (ii) Read Code for IHD; (iii) Read Code for MI; (iv) Read Code for history of angina pectoris; (v) Read Code for coronary artery operations	Search of computer and paper records by 2 observers, validated if found any of: (i) documented ECG change consistent with MI or IHD; (ii) evidence of rise in cardiac enzymes; (iii) positive treadmill, confirming presence of exercise – induced ischaemia;	25 practices selected using stratification by area, urban/rural and partnership size covering 9.4% of population of Northern Ireland. Random sample of 10% of patients found using criteria for CHD ($n = 727$) included in study

TABLE 1 *Continued*

Study	Computer system/software	Coding system	Quality of consultation reporting	Completeness of morbidity register	Correctness of morbidity register	What assessed	'Gold standard'/ assessment criteria	No. of records/ patients/GPs/ practices assessed
de Lusignan <i>et al.</i> (2002) ¹⁷ UK	MediPlus	Read Codes	✓			(i) % of notes linked to diagnosis; (ii) % of notes with more specific Read Code (level 3 or lower); (iii) % of problems/ diagnosis with more specific code. All for first quarter of 2000	(iv) positive coronary angiography; (v) hospital letter confirming coronary artery surgery; (vi) other tests, e.g., thallium scan, dobutamine stress test Whether quarterly feedback improved quality; GPs grouped by year (1992–2000) they joined MediPlus database and effect of time in scheme on quality and whether quality increased over time.	Data from 576 GPs
Jick <i>et al.</i> (1991) ⁹ UK	VAMP	OXMIS		✓	✓	Diagnosis of patients with NSAID prescriptions	GP paper record of hospital discharge letter 1 month before to 14 months after prescription	2491 patients from 58 practices who were first time users of NSAID; 1191 of these had letters forwarded to research group
Jick <i>et al.</i> (1992) ¹⁰ UK	VAMP	OXMIS		✓	✓	Diagnosis of patients with NSAID prescriptions	GP paper record of hospital discharge letter 1 month before to 14 months after prescription	292 patients from further 35 practices who were first time users of NSAID; 126 of these had letters forwarded to research group
Nazareth <i>et al.</i> (1993) ¹¹ London	VAMP	OXMIS/ ICD9	✓	✓	✓	(i) Computer diagnosis (based on ICD9 code) of psychosis. (ii) Extent of recording of consultations	(i) Compared with paper case notes; consensus rating of 3 observers (1 blind to computer category). Diagnosis from case notes was made on basis of criteria in ICD9, the American Psychiatric Association diagnostic manual and a syndrome checklist. Strict criteria definition was positive on all 3 criteria, broader criteria definition was positive on 1 or more. (ii) Paper notes	(i) 13 practices (a) 1 in 2 patients with schizophrenia computer diagnosis ($n = 102$), all patients with affective psychosis diagnosis ($n = 71$) or other psychosis ($n = 78$); (b) all patients with relevant prescription but no computer diagnostic code of psychosis ($n = 194$); (c) random sample of 8000 case notes. (ii) 141 patients from each of 11 practices

Turnbull <i>et al.</i> (1996) ¹² UK	VAMP	OXMIS		✓	Cases of anorexia nervosa and bulimia nervosa from the GPRD	(i) Consensus of 2 psychiatrist specialists in eating disorders viewing complete computer record based on DSM-IV criteria (ii) Psychiatrists' decision and information on symptomatology obtained via questionnaire to patient's GP	(i) 104 randomly selected anorexia nervosa patients and 100 bulimia nervosa. (ii) 25 randomly selected cases of each
Hansell <i>et al.</i> (1999) ¹³ UK	VAMP	OXMIS	✓	✓	GPRD diagnostic prevalence for asthma, COPD, hayfever and tuberculosis for 12 month period (1991/1992)	(i) Prescription prevalence in same period (ii) Prevalence of prescription in same period plus diagnosis at any time	Practices covering 2.9% of the population.
Lawrenson <i>et al.</i> (2000) ¹⁴ UK	VAMP	OXMIS		✓	Cases of venous thromboembolism (VTE) from the GPRD defined as code for deep venous thrombosis or pulmonary embolism with treatment of anticoagulant and prescription for combined oral contraceptive	Evidence from hospital letters and discharge summaries sent by GP	Data from 618 practices let to 277 non-fatal VTE and 8 deaths; details from 169 were obtained from the GPs
Lewis <i>et al.</i> (2002) ¹⁵ UK	VAMP	OXMIS		✓	Cases of ulcerative colitis (based on OXMIS code), Crohn's disease (codes for Crohn's disease, regional colitis, regional enteritis, regional ileitis) and inflammatory bowel disease nos (based on OXMIS code or having code for both ulcerativecolitis and Crohn's disease) from GPRD	(i) Questionnaire sent to GP: case defined as highly probable if GP reported gastroenterology consultation, surgery or intestinal biopsy resulting in diagnosis, probable if based on endoscopy or barium radiography, possible if based solely on GP opinion, probably not if no confirmatory data including GP opinion. (ii) % of previously diagnosed patients receiving appropriate prescription in 1 calendar year	(i) GPs of 170 cases (157 responded with usable information)—49 with Crohn's disease, 88 with ulcerative colitis, 20 with inflammatory bowel disease nos (ii) Crohn's disease 2826 patients, ulcerative colitis 4009 patients

TABLE 1 *Continued*

Study	Computer system/software	Coding system	Quality of consultation reporting	Completeness of morbidity register	Correctness of morbidity register	What assessed	'Gold standard'/ assessment criteria	No. of records/patients/GPs/practices assessed
Margolis <i>et al.</i> (2002) ¹⁶ UK	VAMP	OXMIS		✓	✓	Patients with code for pressure ulcer and no code for chronic wound in 3 months after diagnosis from GPRD	Information from questionnaire sent to patient's GP. Gold standard definition was '... wound that occurred as a localized area of necrosis with a noticeable destruction of skin and deeper soft tissues in an area of bony prominence, and primarily on an individual who was immobile'	2 random samples: 65 patients with code for pressure ulcer, 47 (72%) of GPs replied with information; 65 patients without pressure ulcer, 57 (88%) replied with information

TABLE 2 Results from the studies in the review

Study	Quality of consultation reporting	Completeness of morbidity register	Correctness of morbidity register
Mant and Tulloch (1987) ²⁵		Cancer 52% of patients on register (practice range 45–57%), diabetes 72% (53–89%), MI 43% (37–46%), epilepsy 44%, hypertension 69%, thyroid disease 48%; no 'important' difference between practices using different methods of entry	
Coulter <i>et al.</i> (1989) ²⁴		Epilepsy 49% on register (practice range 30–78%), diabetes 72% (18–97%, 5 practices >90%), thyroid 68% (42–84%), asthma 58% (33–72%); no practice consistently high or low	
Gilliland <i>et al.</i> (1992) ⁷	Study group improved from 15% (surgery and home visit) of consultations given problem/diagnosis to 73% (surgery) and 72% (home visit) over 1 year. Control group's improvement was 26–32% (surgery) and 28–31% (home visit)		
Pringle <i>et al.</i> (1995) ²¹	82% (practice range 67–93%) of diagnostic items in paper notes or computer were on computer. Diagnoses sometimes coded on computer when not seen on videotape by observers. Items missed on computer not of clinical importance	97% of diabetics on computer, glaucoma 92%	No subjects incorrectly recorded as having diabetes or glaucoma; 54% (practice range 46–65%) of asthma patients had received medication, 70% (47–92%) for CHD
Scobie <i>et al.</i> (1995) ²²	91% (12% computer only) of consultations found in computer and/or paper notes were on computer; % of reasons for consulting given code were 74, 80 and 99.5% for 3 practices	Proportion of cases on register: diabetes 5/6 practices >90%, asthma 5/6 practices >70%, angina 2/6 practices >60%; epilepsy 4/6 practices >60%	
Neal <i>et al.</i> (1996) ²⁰	Consultation in notes only ranged from 0.9 to 4.5% by practice; consultations on computer only ranged from 7 to 10%; GPs more likely to fail to record consultations with older females who had more consultations		
Pearson <i>et al.</i> (1996) ²⁷	4516/4685 (96%) records had Read Code. 948/1090 (87%) records checked had appropriate Read Code		
Whitelaw <i>et al.</i> (1996) ¹⁸		Composite sensitivity ranged from 0.4 to 1 over practices (median 0.75). Poor median sensitivity for hypertension, depression, stroke and breast tumour (all <0.6); high for diabetes, epilepsy, glaucoma, gout (>0.8, objective diagnoses); overall diagnostic median sensitivity 0.67	Median PPV 1 across practices for all diagnoses.
Bottomley (1997) ²⁸		Of those practices (9) with CHD register, 2.5% of all patients had nitrates prescribed, 4.3% of all patients were on CHD register	

TABLE 2 *Continued*

Study	Quality of consultation reporting	Completeness of morbidity register	Correctness of morbidity register
Hooker <i>et al.</i> (1999) ⁸		Approximately 70% of hypertensives found in paper notes were on computer	Approximately 70% of hypertensives found on computer were in paper notes
Thiru <i>et al.</i> (1999) ²³		CHD diagnosis prevalence > medication diagnosis prevalence at EMIS practice, similar at Meditel; asthma diagnosis prevalence > medication prevalence at EMIS practice, less at Meditel; glaucoma diagnosis prevalence > medication and procedure prevalence at both practices; diabetes diagnosis prevalence > medication prevalence at Meditel practice, query not run at EMIS practice	
Gray <i>et al.</i> (2000) ¹⁹		Search using G3 and the 5 drugs identified 89/93 (96%) IHD cases; G3 alone found 47%; 5 drugs without G3 found 89% nitrate alone found 47%.	G3 and the 5 drugs = 33% of cases confirmed; G3 alone = 83%; 5 drugs without G3 = 32%, nitrate alone = 83%
Moher <i>et al.</i> (2000) ²⁹		72% of 1979 cases found in both paper notes and register (practice range 29–100%), 27% of cases in notes only	1% on register only—all subsequently found to be definite cases
Hassey <i>et al.</i> (2001) ⁶	98% of 1029 consultations recorded, 12 of missing 20 were DNAs ^a	Mean sensitivity 87%; ratio true positives/false negatives combined with no. of false negatives/10 000 patients for asthma (68), iron deficiency anaemia (32), hypothyroidism (31) and IHD (24) indicate priority areas to identify true undiagnosed cases	True diagnosis in 96% of cases, true absence of diagnosis in 99.5%
Connolly <i>et al.</i> (2002) ²⁶			462/727 (64%) validated. Read Codes alone without nitrate prescription had PPV of 53%
de Lusignan <i>et al.</i> (2002) ¹⁷	Mean % of notes linked to diagnosis ranged from 77 to 92% across groups defined by joining date of MediPlus database; notes with specific Read Code: range 55–77%; problems with specific Read Code: range 71–88%. Only latter appeared to improve significantly with feedback ($P = 0.004$). No sign of joining percentages improving over time		
Jick <i>et al.</i> (1991) ⁹		87% of 1191 had diagnosis on computer record that reflected that in consultant's letter with similar date. Includes appropriate code for 163/190 (86%) of those with newly diagnosed drug-induced disease	
Jick <i>et al.</i> (1992) ¹⁰		96% of 126 had diagnosis on computer record that reflected that in consultant's letter. Includes appropriate code for all 22 with newly diagnosed drug-induced disease	
Nazareth <i>et al.</i> (1993) ¹¹	27% consultations only in paper notes; 25% consultations only on computer	Out of 8000 case notes, 50 psychosis patients identified of which 37 had diagnosis on computer; 18/194 with	65/102 (64%) schizophrenic patients confirmed (strict criteria), 91 (89%) broad criteria, 91%

TABLE 2 *Continued*

Study	Quality of consultation reporting	Completeness of morbidity register	Correctness of morbidity register
Turnbull <i>et al.</i> (1996) ¹²		prescription but no computer diagnosis had no psychosis based on case notes	confirmed as non-organic psychosis. 86% of affective psychosis and 98% of other psychosis patients confirmed non-organic psychosis. PPVs from 8000 sample for computer diagnosis: 71% (schizophrenia-strict), 81% (schizophrenia- broad), 91% (non-organic psychosis-strict) and 100% (non-organic psychosis-broad) Computer records alone: 65/104 anorexia nervosa patients definite; no appropriate data for bulimia nervosa. Computer records and questionnaire: 15/25 anorexia nervosa met DSM-IV criteria, 3 likely; 13/25 bulimia nervosa definite, 3 likely;
Hansell <i>et al.</i> (1999) ¹³		Asthma and COPD: current diagnosis prevalence alone much lower than prescription rates, particularly for under 16 (asthma) and over 45 (both) age groups, and lower but closer to prescription plus diagnosis ever rates; hayfever: current diagnosis prevalence alone much lower than prescription rates and lower but closer to prescription plus diagnosis ever rates. Tuberculosis: current diagnosis prevalence alone much lower than prescription rates but slightly higher than prescription plus diagnosis ever rates	
Lawrenson <i>et al.</i> (2000) ¹⁴			167/169 (99%) had confirmation admitted to hospital from details sent by GP. Diagnosis of VTE supported by hospital investigations in 141/169 (83%)
Lewis <i>et al.</i> (2002) ¹⁵			Crohn's disease 42 (86%) highly probable have inflammatory bowel disease, 3 (6%) probably not; ulcerative colitis 77 (88%) highly probable, 3 (3%) probably not; inflammatory bowel disease nos 15 (75%) highly probable, 4 (20%) probably not; total 134 (85%, 95% CI 79%, 90%) highly probable, 10 (6%) probably not. 48% (Crohn's disease) and 60% (ulcerative colitis) of patients had at least one relevant prescription
Margolis <i>et al.</i> (2002) ¹⁶		3/57 patients without pressure ulcer code had ulcer based on GP questionnaire.	All 47 patients with pressure ulcer validated by GP questionnaire

DNA = did not attend.

Several studies only included the highest performing practices in terms of computer use, or those which had undergone extensive training in the use of computers and coding.^{6,18,21,23} To be included in the GPRD also requires a period of training. Problems were sometimes encountered enrolling practices into the study, due either to the practice lacking the information required (e.g. a morbidity register) or to the practice refusing to enter the study. For example, in one study, 31 out of 38 practices approached were unable to supply the required data.²⁰ In another, only nine out of 48 practices contacted supplied useable information for the study.²⁸

Quality of consultation recording

Eight studies looked at the completeness of consultation recording.^{6,7,11,17,20–22,27} This was in terms of either the extent to which consultations were being recorded on the computer^{6,11,20,22} or the extent to which morbidity codes were being allocated to recorded consultations.^{7,17,21,22,27} The former were all compared with paper notes, except one study which used an appointment book.⁶

The extent of electronic consultation recording was generally high (>90%), except in one study where only three-quarters of consultations were on computer.¹¹ In this study, a quarter of consultations were not in the paper notes. In other studies, ~10% of consultations were on computer but not in the paper notes.^{20,22}

In terms of morbidity codes allocated to consultations, there was high variability between the practices studied, even within the separate studies, ranging from 67 to 99%.^{7,21,22,27} This may have improved since these studies were published (1992–1996). de Lusignan *et al.* suggested that feedback improved the detail of coding but, in general, in their study, ~80% of problems had Read Codes down to level 3 or lower (i.e. were more specific).¹⁷

GPs in one study suggested that mental and psychological problems were often not recorded due to their difficulty of coding and that chronic problems were also under-recorded.²² In another study, doctors were more likely to fail to record consultations with older females who were more frequent consulters.²⁰ Discrepancies were considered to be due to unavailability of a computer (e.g. for a home visit), lack of motivation or forgetfulness. One study suggested that simple requests to GPs to record diagnoses, with or without feedback of previous results, could greatly improve the completeness.⁷

The only study which assessed consultation content as a measure of comprehensiveness of the codes given for the contact did so by reference to videotapes of consultations.²¹ The authors suggested that items missed were not of clinical importance. Diagnoses not apparent from observation of the videotape were sometimes coded on the computer.

Another study suggested that 87% of 1090 records checked had the appropriate Read Codes when

compared with paper notes, although little information is given on the methodology for assessing this 'appropriateness'.²⁷

Morbidity registers

Heart disease registers, including angina, myocardial infarction (MI) and ischaemic heart disease (IHD), have been assessed most often (10 studies). These have been by comparison with prescribed medication (typically nitrates),^{21–23,28} combinations of information (including drugs and hospital letters) in paper notes,¹⁸ medication and procedures stored electronically,⁶ tests and procedures^{19,26,29} and hospital discharge information collected by the Health Authority.²⁵

Completeness of heart disease registers appears poor. Seventy-two percent of patients with validated coronary heart disease (CHD) based on related information held in paper notes and computer records were on the registers of 18 practices (one practice's register was not computerized).²⁹ One study reported that only two out of six practices had >60% of those with angina medication with an angina code,²² and another study reported that a search using the IHD Read Code (G3) identified only 47% of probable IHD patients.¹⁹ Forty-three percent of subjects identified as having left hospital following a MI were coded at four practices.²⁵ Hassey *et al.* in their single practice study noted IHD as a high priority area for identifying undiagnosed cases (an estimated 24 in a population of 10 000), although 96% of their suspected cases, based on medication and procedures, were identified by Read Code.⁶ Whitelaw *et al.* detected a median 60% of suspected angina cases and 80% of MI cases in the electronic records of 41 practices.¹⁸ However, two studies comparing the prevalence of CHD diagnostic coding with that of nitrate medication at the same practices reported higher prevalences for the diagnosis.^{23,28} This may suggest that estimation of the prevalence of CHD is improved when based on diagnosis, e.g. not all patients may be on medication, but may also suggest that some patients are incorrectly placed on the register.

Correctness of heart disease coding appears slightly better. Seventy percent of patients (ranging from 47 to 92% between practices) with the morbidity code had appropriate medication in one study.²¹ This figure improves to 83–100%^{6,18,19,29} when the comparison included tests and procedures, although Connolly *et al.*²⁶ could only validate 53% of cases with a CHD Read Code but no nitrate prescription. Moher *et al.* found that surgery contact was the main predictor of being on the CHD register [odds ratio (OR) 2.1]. Other significant predictors were repeat prescriptions (OR 1.6), MI diagnosis (OR 1.5) and revascularization procedure (OR 1.5). Diagnosis of angina, year of diagnosis, age and gender were not significant predictors.²⁹

Seven studies examined completeness of diabetes registers, and three of these also looked at correctness.^{6,18,21} One study compared medication and

diagnostic code prevalence.²³ Gold standards were medication^{22–24}, combinations of related information (electronic and/or non-electronic)^{6,18,21} and hospital discharge information.²⁵ Completeness of diabetes registers appears high, with consensus of 90% or more of diabetics identified in practice. The exception to this used a comparison with hospital discharge data (72%) and was the oldest study reviewed (data collected 1982–1984).²⁵ Correctness was agreed to be near to maximum.

Asthma and COPD registers have been investigated in seven studies, four of which examined completeness^{6,18,22,24} and three of which examined correctness.^{6,18,21} These were in relation to medication^{6,21,22,24} and paper information including medication.¹⁸ Two studies compared medication and diagnostic prevalence.^{13,23} Quality of the asthma registers was variable: compared with medication, five out of six practices coded 70% of patients on asthma medication,²² but only 58% were coded in Coulter's older study.²⁴ Asthma was again highlighted by Hassey *et al.*⁶ as a high priority area for identifying undiagnosed cases. They estimated 68 unrecorded cases in a population of 10 000. Pringle *et al.* found only 54% of asthma-coded cases in four practices with the appropriate medication.²¹ Comparison of medication and diagnostic prevalences gave variable results.^{13,23}

Completeness of epilepsy coding was moderate (generally 40–70%) when compared with medication^{22,24} and hospital discharge information,²⁵ but higher in two other studies (>90%) using medication⁶ or combinations of relevant paper-held evidence¹⁸ as the gold standard. These also found high correctness.

The completeness and correctness of glaucoma coding has been found to be >80%^{6,18,21} whilst the prevalence of glaucoma based on diagnostic coding was greater than for that based on medication and procedures at two practices.²³

The completeness of thyroid conditions ranged from 42 to 84% across practices in the two oldest studies.^{24,25} Hassey *et al.*'s study found high correctness and completeness of 82% for hypothyroidism and 98% for hyperthyroidism.⁶

The evidence for hypertension is mixed. Hassey *et al.* found high completeness (98%) and correctness for hypertension when compared with medication.⁶ However, Whitelaw *et al.* found poor completeness in their study (median across 41 Scottish practices of 43%).¹⁸ One other study found correctness and completeness for hypertension coding of ~70% in a random sample of 100 elderly patients in one practice compared with paper notes, but gave little information on methodology.⁸

The completeness and correctness of gout has been found to be good.^{6,18} However, the results for cancer have been variable. Hassey *et al.* in a triangulation with drugs and other treatments found high completeness and correctness for breast and

prostate cancer, although both had <60 cases.⁶ Whitelaw *et al.*, however, found poor completeness for breast tumour (median across practices of 57%).¹⁸ Mant and Tulloch also found poor completeness (52%) for cancer compared with hospital discharge information in their older study.²⁵

Whitelaw *et al.* found poor completeness for depression (47%).¹⁸ A study of GPRD London practices assessed diagnoses of schizophrenia, affective psychoses and other psychoses compared with information contained in paper notes.¹¹ Psychosis diagnosis appeared to be ~74% complete based on just 50 cases identified through a random search of paper notes. Correctness was high for 149 affective and other psychosis computer-coded patients, although slightly lower for 102 schizophrenia-coded patients (64–89%) depending on strictness of criteria for diagnosis. However, >90% of the schizophrenia-coded patients were deemed to have a non-organic psychosis.

Hassey *et al.* found good completeness and correctness for rheumatoid arthritis in their practice,⁶ but completeness was poorer across 35 Scottish practices (median 67%).¹⁸

Other conditions have been assessed in the GPRD studies. Lawrenson *et al.* explored cases of venous thromboembolism (VTE), defined by morbidity code with evidence of treatment. They compared this with hospital investigation information forwarded to the research team by the GP. VTE was supported in 83% of the 169 cases.¹⁴ In a further study, a correctness of 85% of bowel disease diagnosis was found out of 157 cases based on a valid reason for diagnosis (e.g. intestinal biopsy, gastroenterology consultation, surgery). This information was obtained by a questionnaire to the GP asking how the diagnosis was made. However, only around a half of almost 7000 inflammatory bowel disease patients had a relevant prescription.¹⁵ One study suggested high correctness and completeness of pressure ulcer morbidity coding based on information from a questionnaire completed by the GP;¹⁶ another study suggested correct diagnosis of ~60% in anorexia nervosa and bulimia coded cases when experts assessed information on the practice computer record and a GP questionnaire on symptoms and treatment.¹² However, both studies suffered from small sample sizes.

Finally, two other GPRD studies suggested that ~90% of patients prescribed non-steroidal anti-inflammatory drugs (NSAIDs) and with a hospital consultant's letter forwarded by their GP had a diagnosis on computer that matched the diagnosis in the letter.^{9,10}

Discussion

The quality of morbidity coding appears variable. Conditions with clear diagnostic features such as

diabetes have higher quality recording than conditions with more subjective criteria such as asthma. It is difficult to ascertain an improvement in quality over time. Although, two studies in the 1980s looking at chronic disease registration in Oxford practices found poor completeness of registers,^{24,25} quality has tended to be mixed in studies performed both in the 1990s and since the start of 2000. Hassey's more recent study (2001) in a trained practice found reasonably high completeness and correctness but suggested a number of morbidities where some cases may not have the relevant morbidity code recorded.⁶

The GPRD studies have shown reasonable correctness and completeness of morbidity registers (although poorer for some diseases such as anorexia nervosa and bulimia). Of the other multipractice studies, the studies on the GPASS group in Scotland,¹⁸ the MediPlus database¹⁷ and practices in Northern Ireland²⁶ all suggest there is room for improvement in the areas assessed, but showed variation across practices. Most studies were of a reasonable size, although for rarer diseases the number of cases were small.

The completeness of consultation recording was generally high. However, although most contacts with a GP may have a computer-recorded morbidity code, this does not mean that all problems addressed during a multiproblem encounter are recorded. Only one study²¹ addressed this issue, which requires further research. It may be that only new problems or problems regarded as the most important are recorded, leaving ongoing or minor problems unrecorded.

Completeness and correctness of data entry may rely on the enthusiasm of practices and of individual GPs. GPs may have personal preferences for certain codes which may not always be appropriate. Morbidity coding is subjective and relies on the characteristics and idiosyncrasies of individual GPs.³⁰ Many of the studies reported here looked at practices with explicit interests in recording information electronically or with a substantial amount of training in morbidity coding. Several multipractice studies had to discard practices from their study which were unable to provide suitable data. This biases the sample in favour of the better recorders, leading to a higher quality of recording than that which would be achieved by examining all practices. The majority of studies were also based in one localized area (exceptions include the studies based on the GPRD), which makes generalization difficult.

Evaluation of the accuracy of electronic records is enhanced by use of common coding systems. One question is at what level of a coding system should validation be demonstrable. For example, recording diagnoses only at diagnostic Chapter heading (level 1 of the Read Code) may improve the accuracy of recorded data but is too general for most research or clinical purposes. It could be argued that the insistence on use of

a coding system loses the richness of the material contained in consultations, which may be better expressed in free text descriptions. GPs may feel pressured into using the codes available even if they are inappropriate for that patient.

A variety of gold standards have been used, and completeness and correctness can only be inferred in relation to the quality of the gold standard(s) used. A number of studies used the terms 'sensitivity' and 'positive predictive value'. In diagnostic testing, these terms are used to imply whether patients really do have the disease. To avoid this implication, we have used the terms 'completeness' and 'correctness', as we have only examined internal validity, e.g. where one source of information (e.g. prescription data or paper notes) has been used to justify the existence of a morbidity code.³¹ We have not examined external validity which relates more to the normal interpretations of sensitivity and PPV.

Information in paper notes and medication were most commonly used to validate morbidity coding. Paper notes have to be scrutinized by hand, preferably by two researchers, to improve the reliability and completeness of the paper trawl. A disadvantage is the assumption that the paper notes are accurate. Entries in paper notes may be missing or illegible, and the notes may be poorly organized, which means that searching for information can be difficult. Increases in the number of paperless practices will also mean that, increasingly, only other electronic data can be used as a gold standard. Problems of comparing diagnosis with medication, which may be in paper notes or stored electronically, is that the same medication may be prescribed for many conditions, or diagnoses may be old and patients may no longer need, and therefore have prescribed, a medication. This may be one of the reasons asthma had poorer quality of recording. It also assumes that the medication information is complete and correct. Studies can only include morbidities where a medication is specific to that disease and results cannot be confounded by patients obtaining the medication, e.g. over the counter. Whilst prescriptions, hospital tests and procedures are likely to be the best internal validation for morbidities (as recommended previously⁵), this limits the number that can be validated.

Our review has been more specific in focus than that of Thiru *et al.*⁵ who reviewed methods of assessing the quality of all data contained in electronic records, using any gold standard and including non-UK studies. This has allowed us to give a specific assessment of the quality of computerized morbidity recording in primary care in the UK and to critique methods for assessing this quality.

The decision to include only UK studies in this review will have reduced the number of studies. However, the difference in the structure of primary care services and coding and computer systems between nations makes it difficult to combine studies from different countries.

Estimates based on our search suggest that ~6–10 non-UK studies would have been included although they are unlikely to have uncovered any different methodologies. We have further excluded studies which have compared patient's self-report of disease with that contained in computerized medical records. This would be likely to entail a systematic review of its own, and the objective of such studies may be different. It is difficult to know quite which is being validated, the self-report or the medical records. Studies comparing prevalence rates with external data (from national sources, other practices or at different time points) were also not accepted into the review. Illness prevalences from different localities or at different time points are subject to a range of influences other than the quality of coding, notably variations in prevalence, since certain morbidities or morbidity codes may become more fashionable over time or criteria for a diagnosis may change.³² We also did not search for grey literature and other unpublished studies. It is possible these may have added to our review.

This paper has reviewed the quality of morbidity coding in general practice electronic records. As this review has highlighted, it does appear possible to assess and verify the quality of coding of certain morbidities, and some of these have been shown to be generally well coded; others have poorer levels of coding. Training of practices, as shown in the GPRD studies and in the study by Hassey *et al.*,⁶ can lead to a reasonable quality of coding. As practices increasingly use computers to record consultations and other medical information, there is a need to ensure that there is a high level of completeness and correctness of not just morbidity codes relevant to the consultation, but also information from external sources such as hospital letters. The focus should be now be on methods to encourage and improve the quality of this coding in general practice.

References

- 1 RCGP. *General Practice Computerisation*. London: Royal College of General Practitioners; 2000.
- 2 Mitchell E, Sullivan F. A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980–97. *Br Med J* 2001; **322**: 279–282.
- 3 NHS Information Authority. *The Clinical Terms Version 3 (The Read Codes)*. Birmingham: NHS Information Authority; 2000.
- 4 Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. *J Am Med Inform Assoc* 1997; **4**: 342–355.
- 5 Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. *Br Med J* 2003; **326**: 1070–1072.
- 6 Hassey A, Gerrett D, Wilson A. A survey of validity and utility of electronic patient generated records in a general practice. *Br Med J* 2001; **322**: 1401–1405.
- 7 Gilliland AEW, Mills KA, Steele K. General practitioner records on computer—handle with care. *Fam Pract* 1992; **9**: 441–450.
- 8 Hooker RC, Cowap N, Newson R, Freeman GK. Better by half: hypertension in the elderly and the 'rule of halves': a primary care audit of the clinical computer record as a springboard to improving care. *Fam Pract* 1999; **16**: 123–128.
- 9 Jick H, Jick SS, Derby LE. Validation of information recorded on a general practitioner based computerised data resource in the United Kingdom. *Br Med J* 1991; **302**: 766–768.
- 10 Jick H, Terris BZ, Derby LE, Jick SS. Further validation of information recorded on a general practitioner based computerized data resource in the United Kingdom. *Pharmacoepidemiol Drug Saf* 1992; **1**: 347–349.
- 11 Nazareth I, King M, Haines A, Rangel L, Myers S. Accuracy of diagnosis of psychosis on general practice computer system. *Br Med J* 1993; **307**: 32–34.
- 12 Turnbull S, Ward A, Treasure J, Jick H, Derby L. The demand for eating disorder care—an epidemiological study using the general practice research database. *Br J Psychiatry* 1996; **169**: 705–712.
- 13 Hansell A, Hollowell J, Nichols T, McNiece R, Strachan D. Use of the General Practice Research Database (GPRD) for respiratory epidemiology: a comparison with the 4th Morbidity Survey in General Practice (MSGP4). *Thorax* 1999; **54**: 413–419.
- 14 Lawrenson R, Todd JC, Leydon GM, Williams TJ, Farmer RD. Validation of the diagnosis of venous thromboembolism in general practice database studies. *Br J Clin Pharmacol* 2000; **49**: 591–596.
- 15 Lewis JD, Brensinger C, Bilker WB, Strom BL. Validity and completeness of the General Practice Research Database for studies of inflammatory bowel disease. *Pharmacoepidemiol Drug Saf* 2002; **11**: 211–218.
- 16 Margolis DJ, Bilker W, Knauss J, Baumgarten M, Strom BL. The incidence and prevalence of pressure ulcers among elderly patients in general medical practice. *Ann Epidemiol* 2002; **12**: 321–325.
- 17 de Lusignan S, Stephens PN, Adal N, Majeed A. Does feedback improve the quality of computerized medical records in primary care? *J Am Med Inform Assoc* 2002; **9**: 395–401.
- 18 Whitelaw FG, Nevin SL, Milne RM, Taylor RJ, Taylor MW, Watt AH. Completeness and accuracy of morbidity and repeat prescribing records held on general practice computers in Scotland. *Br J Gen Pract* 1996; **46**: 181–186.
- 19 Gray J, Majeed A, Kerry S, Rowlands G. Identifying patients with ischaemic heart disease in general practice: cross sectional study of paper and computerised medical records. *Br Med J* 2000; **321**: 548–550.
- 20 Neal RD, Heywood PL, Morley S. Real world data—retrieval and validation of consultation data from four general practices. *Fam Pract* 1996; **13**: 455–461.
- 21 Pringle M, Ward P, Chilvers C. Assessment of the completeness and accuracy of computer medical records in four practices committed to recording data on computer. *Br J Gen Pract* 1995; **45**: 537–541.
- 22 Scobie S, Basnett I, McCartney P. Can general practice data be used for needs assessment and health care planning in an inner-London district? *J Public Health Med* 1995; **17**: 475–483.
- 23 Thiru K, de Lusignan S, Hague N. Have the completeness and accuracy of computer medical records in general practice improved in the last five years? The report of a two-practice pilot study. *Health Informatics J* 1999; **5**: 224–232.
- 24 Coulter A, Brown S, Daniels A. Computer held chronic disease registers in general practice: a validation study. *J Epidemiol Community Health* 1989; **43**: 25–28.
- 25 Mant D, Tulloch A. Completeness of chronic disease registration in general practice. *Br Med J* 1987; **294**: 223–224.
- 26 Connolly P, Cupples ME, Cuene-Grandidier H, Johnston D, Passmore P. The importance of validating the diagnosis of coronary heart disease when measuring secondary prevention: a cross-sectional study in general practice. *Pharmacoepidemiol Drug Saf* 2002; **11**: 311–317.
- 27 Pearson N, O'Brien J, Thomas H, Ewings P, Gallier L, Bussey A. Collecting morbidity data in general practice: the Somerset morbidity project. *Br Med J* 1996; **312**: 1517–1520.

- ²⁸ Bottomley A. Methodology for assessing the prevalence of angina in primary care using practice based information in northern England. *J Epidemiol Community Health* 1997; **51**: 87–89.
- ²⁹ Moher M, Yudkin P, Turner R, Schofield T, Mant D. An assessment of morbidity registers for coronary heart disease in primary care. *Br J Gen Pract* 2000; **50**: 706–709.
- ³⁰ Fleming DM. The measurement of morbidity in general practice. *J Epidemiol Community Health* 1991; **45**: 180–183.
- ³¹ Newcombe RG, Altman DG, Bryant TN. Electronic patient records in general practice. Methods of evaluation of electronic patient records entail dangers. *Br Med J* 2001; **323**: 1184–1185.
- ³² Anderson JE. Reliability of morbidity data in family practice. *J Fam Pract* 1980; **10**: 677–683.