Assessment of the completeness and accuracy of computer medical records in four practices committed to recording data on computer

MIKE PRINGLE

PAT WARD

CLAIR CHILVERS

SUMMARY
Background. General practice computer databases are being increasingly seen as a source of data for public health monitoring and commissioning. Such ambitions depend on routine clinical data being recorded with acceptable completeness and accuracy.
Aim. The aim of this study was to assess the completeness and accuracy of the computer medical records in four high-recording general practices.
Method. Four general practices in the Trent Region that use the EMIS computer system, and were known to be high recorders of clinical data on their computer databases, were selected. A retrospective analysis of the computer records, a prospective comparison of a sample of computer records with manual records, and a prospective comparison between videorecorded consultations and their manual and computer records were undertaken.
Results. Checks for completeness in computer recording of diabetes mellitus and glaucoma showed high levels of accurate recording, 97% and 92% respectively. Prevalence rates between practices were reasonably comparable. No practice consistently, across 10 diagnoses, recorded prevalences higher or lower than the other practices; those diagnoses with recognized objective diagnostic criteria were recorded with a more consistent prevalence than those without. Lifestyle data recording was low; overall, smoking habits and alcohol consumption were recorded for 52% and 38% of patients aged over 16 years, respectively. Comparison of the manual records with the computer records showed that the computer records were sufficiently complete with regard to diagnoses (82% of all items recorded), prescriptions (100%) and referrals (67%), but missed most of the remaining data that a manual record captured. The videorecorded validation study showed that there were no important lapses in the recording of diagnoses, prescriptions or referrals when the computer recording was compared to the actual process of the consultations.
Conclusion. In these four high-recording practices the data in computer records were of sufficient completeness and accuracy to allow meaningful data aggregation for some diagnoses, prescriptions and referrals. Standardized protocols for defining which patients are included and excluded from major disease groups are required.

M Pringle, FRCGP, professor, Department of General Practice, Nottingham University Medical School. P Ward, research assistant and C Chilvers, HonMFPHM, professor of epidemiology, Department of Public Health Medicine and Epidemiology, Nottingham University Medical School.

Introduction

THERE is increasing interest in the use of computer databases in general practice.[1-5] Commissioners of health care, traditionally limited to mortality data and secondary care process measures, wish to have timely data on the socioeconomic and lifestyle characteristics, morbidity, resource usage and anticipated needs of their populations.[6] Monitoring of many *Health of the nation* targets requires lifestyle data for populations.[7] General practices are beginning to construct business plans based on an assessment of the health needs of their populations, in order to bid for and rationalize the use of health service funding. Fundholding practices, in particular, are moving from contracting to purchasing and will eventually move into commissioning. As they do so their information needs will become more sophisticated.

Some problems in aggregating practice data are inherently technical: communication links must allow the uncorrupted transmission of patient data in an anonymous, confidential manner, and a common coding structure for diseases, activities and patient characteristics must be established to ensure accurate inter-practice comparison. These problems are now largely solved by the use of Read codes for recording data,[8] *MIQUEST* for the interrogation of practice databases by a third party without patient confidentiality being breached,[9] and *EDIFACT* standards for communicating between systems.[10] The remaining problems are largely focused on the data themselves: clinical databases are often incomplete, inaccurate and inconsistent.

From the few studies that have been published on the validation of general practice computer databases it is clear that recording of acute illness on general practice computers may occur for only a quarter of the total prevalence[3] but that doctors are capable of recording high-quality data for some chronic diseases.[11,12] Lifestyle data are poorly recorded on computers and socioeconomic data are virtually absent on United Kingdom general practice computers.[13] Some practices have accepted low cost computer systems[14] without accepting the need for rigour in recording.[2,3]

This study was designed as part of an evaluation of the scope for the aggregation of databases from general practices and addresses the need to validate the computer systems of four high-recording practices for completeness and accuracy.

Method

Four practices in the Trent Region that use the EMIS general practice computer system were recruited. All four were selected from a list, supplied by the computer system company, that identified the practices with the highest commitment to recording data on computer. This study therefore examines the quality of the data in general practices that were likely to be among the best in terms of data recording. It is not a study of the data quality in average general practices but a measure of what is being achieved in a small number of high recorders.

The study lasted 18 months, from July 1992 to January 1994. Each practice's database was examined for recorded diagnoses, lifestyle data and activity such as prescribing and referrals. Specific validation checks were designed and were applied to each practice's data by P W.

Consistency of computer recorded prevalence of 10 diagnoses

The computer databases were retrospectively examined for the recording of 10 major diagnoses in order to look for variations in recorded prevalence that might indicate incomplete recording. The 10 diagnoses were asthma, coronary heart disease, dementia, depression, diabetes mellitus, glaucoma, hypertension, multiple sclerosis, osteoarthritis of the knee, and stroke (having had a stroke ever or in the previous two years). Data from the four practices were compared with data from the fourth national morbidity study.[15]

Accuracy of computer recording of four diagnoses

In all four practices the computer recording of four major chronic diagnoses was retrospectively assessed and compared with the prevalence of the four diagnoses found from manual medical records and the repeat prescribing system (all four practices used the EMIS integral repeat prescribing system) or found by cross-checking in the computer records for other related diagnoses. The diagnoses used for this diagnostic validation within each practice were: diabetes mellitus, glaucoma, asthma and coronary heart disease. The first two were chosen to assess cases that were not recorded on computer or were inappropriately coded and the last two were chosen to allow examination of the ratio of active disease to overall recording of disease.

The recording of diabetes mellitus was checked in each practice for completeness and accuracy by comparing the computer diagnostic register with the repeat prescribing system for all glucosuria and blood glucose diagnostic tests and for all hypoglycaemic agents, including insulin. Where discrepancies were found, the continuation cards, problem lists and prescribing details in the manual medical records were consulted. The manual medical records for patients with a computer recorded diagnosis of diabetes mellitus but no related prescription were checked. A similar assessment was made for glaucoma by cross-checking against computer records for trabeculectomy and by checking for evidence of recorded diagnoses in the manual medical records.

In each practice, records on the computer diagnostic register for asthma were compared with a list of patients who had received any inhaled bronchodilator or inhaled steroid treatment in the previous two years. For coronary heart disease, the computer diagnostic register was compared with a list of patients having had a prescription for nitrates in the previous two years.

Completeness of lifestyle and socioeconomic data

The extent of computer recording of lifestyle data (smoking habits and alcohol consumption) and of socioeconomic data (social class, occupation and ethnicity) were analysed.

Completeness of computer recording of consultations

To assess the completeness of consultation recording in the computer databases, 50 consecutive consultations for each doctor in each of the four practices were selected for study, starting at a random time in the week. The practices were visited by P W in the week following the recording of the consultations and the computer records and manual records of the consultations were examined to compare the content of the records. This was done by comparing the consultation content 'items', recorded as sep-

arate entries, in the computer record with those in the manual record.

One of the four practices was then recruited for the second phase of this prospective assessment in order to assess how much information from consultations was not recorded in the computer records or manual records. Fifty consecutive consultations by each of four general practitioners with consenting patients were videorecorded with the general practitioner's knowledge; the videorecording and the computer and manual records were examined for their content. All patients to be videorecorded were asked for prior signed informed consent and ethics committee approval was obtained.

For the videorecording, the following definition of a 'topic' was used: a topic exists when both parties use at least one phrase or sentence in its discussion; or if one party discusses it and takes an appropriate action (for example, the general practitioner measures the patient's blood pressure); or if a prescription review takes place without mentioning the underlying topic explicitly. A topic was only recorded once in each consultation even if mentioned several times. M P and P W analysed a pilot sample of videorecorded consultations to develop the methodology applied in this study and only when intra-rater and inter-rater kappa coefficients[16] of more than 0.8 were obtained for number and type of topics discussed, and for consultation content items, was the main analysis undertaken.

Results

Consistency of computer recorded prevalence of 10 diagnoses

Table 1 shows the prevalence of 10 diagnoses in the four practices and allows assessment of inter-practice variation and comparison with the fourth national morbidity study data.[15] For the 10 diagnoses checked, no practice consistently, across all diagnoses, recorded prevalences higher or lower than the other practices.

Where the diagnosis has recognized objective diagnostic criteria (for example in diabetes mellitus, glaucoma, multiple sclerosis, osteoarthritis of the knee, and stroke) the prevalence rates were consistent between the four practices.

The more subjective diagnoses were less consistently recorded. The prevalence of hypertension, for example, ranged between 4.4% and 7.9%. By contrast, asthma showed a similar prevalence across the practices. The consistency of recording of coronary heart disease lay between those of hypertension and asthma. The recorded prevalence of dementia ranged from 0% to 0.4%, but practice C reported that it had not been placing emphasis on this diagnosis. Depression offered a four-fold variation in recorded prevalence, suggesting that it was inconsistently defined or recorded.

Accuracy of computer recording of four diagnoses

Results of the assessment of the four diagnoses in the four practices are shown in Table 2. In all the practices the computer diagnostic registers underestimated the prevalence of diabetes mellitus, but no diagnoses incorrectly recorded as diabetes mellitus were found. Overall, 785 of 812 confirmed diabetics (96.7%) were correctly identified from the computer diagnostic registers alone. In two practices (B and C) the recording of glaucoma was apparently complete; in practices A and D a few cases of glaucoma would have been missed on a diagnostic search if the computer records had not been cross-checked against trabeculectomy, 17 of 222 cases in the four practices (7.7%). No diagnosis was incorrectly recorded as glaucoma.

Overall in the four practices, 53.6% of the asthma patients on the computer diagnostic registers had received medication for asthma in the previous two years. The ratio of those who had

Table 1. Prevalence of 10 diagnoses, using computer diagnostic registers in four practices, and comparison with national prevalence data.[a]

| Diagnosis | % of practice population with diagnosis in practice | | | | | |
	A (n = 5575)	B (n = 11 687)	C (n = 9896)	D (n = 10 297)	A – D (n = 37 455)	National data[a,b]
Asthma	6.8	6.6	6.2	6.7	6.5	9.1
Coronary heart disease	4.1	3.5	2.6	2.1	3.0	4.3
Dementia	0.4	0.1	0	0.1	0.1	0.4
Depression	3.5	10.6	3.8	2.5	5.5	–
Diabetes mellitus	1.9	2.6	1.7	2.0	2.1	2.7
Glaucoma	0.5	0.5	0.4	0.7	0.5	0.3
Hypertension	6.0	7.9	4.6	4.4	5.8	10.3
Multiple sclerosis	0.3	0.2	0.1	0.2	0.2	0.2
Osteoarthritis of knee	2.0	1.7	1.8	0.9	1.6	–
Stroke						
Ever	1.3	0.8	1.1	1.1	1.1	1.3
In last 2 years	0.4	0.1	0.2	0.4	0.3	–

n = number of patients in practice.[a]Consultation rates per 100 person years at risk, from the fourth national morbidity study.[15] [b]Data not available for depression, osteoarthritis of knee, and stroke in the last two years.

received treatment in the previous two years to those who had ever been diagnosed as asthmatic, as recorded on the computers, appeared reasonably consistent between practices, from 45.6% to 65.3%. The ratio of the number of patients issued a prescription for nitrates in the previous two years to those ever diagnosed with coronary heart disease, as recorded on the computers, ranged between practices from 47.4% to 91.5%.

Completeness of lifestyle and socioeconomic data

All the practices combined had recorded the smoking habits of 52.1% of patients aged over 16 years, ranging between practices from 46.1% to 55.8%. Alcohol consumption was recorded for 37.5% of patients aged over 16 years, ranging between practices from 27.1% to 71.4%. Occupation was not routinely recorded by any of the practices, and in practice A (which had the highest recording for this item) only 165 occupations were recorded on the computer. Social class and ethnicity had not been recorded for any of the patients in the study.

Completeness of computer recording of consultations

Information in either manual or computer records or in both were examined for 50 consecutive consultations by each doctor in the

four practices. In total, practice A recorded 245 items in 200 consultations (1.2 items per consultation), practice B 331 items in 250 consultations (1.3), practice C 285 items in 250 consultations (1.1), and practice D 334 items in 300 consultations (1.1). An analysis of the content of these items is shown in Table 3. The manual records offered a more complete picture of the consultation content in all variables except diagnosis and prescriptions, for which the computer records were, overall, more complete. For referrals, the computer recording was comparable to the manual recording in two practices, but in two practices (B and C) fewer than half of the referrals identified were recorded on the computer. In two practices (A and C) fewer than a third of investigations were recorded on the computer. Wide inter-practice variations in recording rates were seen for some topics.

Practice A was recruited to undertake the more detailed validation of consultation recording, using videorecordings (Table 4). Compared with the videorecordings, both manual records and computer entries failed to capture some data. An examination of those videorecorded topics missed in the manual or computer records of the same consultations revealed that they were predominantly those to which a medical problem was not attached when the videorecording was coded; they were often social or

Table 2. Records found on computer diagnostic register search for four diagnoses and additional cases found on checking for completeness of recording or number of recorded diagnoses found to be cases of active disease, in four practices.

| Diagnosis | No. of records of diagnosis in practice | | | |
	A	B	C	D
Diabetes mellitus				
Computer diagnostic register	104	303	170	208
Additional cases found from manual records/repeat prescribing system	1	13	10	3
Glaucoma				
Computer diagnostic register	27	63	39	76
Additional cases found from manual records/cross-check against trabeculectomy	7	0	0	10
Asthma				
Computer diagnostic register	377	766	610	694
Asthma medication in last 2 years (% of computer recorded cases)	246 (*65.3*)	349 (*45.6*)	342 (*56.1*)	374 (*53.9*)
Coronary heart disease				
Computer diagnostic register	226	405	258	217
Medication with nitrates in last 2 years (% of computer recorded cases)	199 (*88.1*)	192 (*47.4*)	236 (*91.5*)	143 (*65.9*)

Table 3. Occurrence of items in 50 consultations per doctor in four practices (1000 consultations) in manual and computer records as a percentage of total occurrences found in either source.

Consultation item	% of items (range over four practices) recorded in records	
	Manual	Computer
Symptoms (n = 966)	96.8 (90.2–100)	37.0 (5.9–95.7)
Diagnosis (n = 828)	65.5 (41.8–82.4)	81.8 (67.2–92.9)
Prescription (n = 696)	87.2 (85.1–89.3)	100
Examination (n = 467)	95.7 (87.7–99.1)	27.2 (2.3–87.7)
Site of symptoms (n = 383)	90.9 (83.8–99.1)	40.5 (5.3–90.0)
Review arranged (n = 330)	78.8 (78.6–98.8)	42.7 (7.4–81.1)
Numerical finding (n = 320)[a]	91.6 (87.6–95.2)	60.6 (23.8–89.5)
Duration of symptoms (n = 242)	96.3 (85.9–100)	20.7 (0–75.0)
Preventive advice (n = 203)	82.3 (66.3–95.0)	43.8 (6.1–85.5)
Investigation arranged (n = 146)	94.5 (90.0–100)	54.8 (20.0–92.5)
Previous test results (n = 103)	89.3 (82.6–94.7)	57.3 (21.1–92.5)
Referral (n = 45)	84.4 (71.4–100)	66.7 (30.0–92.9)
Preventive action (n = 23)[b]	91.3 (50.0–100)	56.5 (0–100)
Treatment given (n = 16)	81.3 (75.0–100)	81.3 (62.5–100)

n = number of occurrences of item recorded in either manual or computer records. [a]For example, blood pressure. [b]For example, immunization.

minor medical discussions. No instance of an important clinical problem eluding the manual or computer records was seen. There was consistency in the frequency of recording of items in the manual records and on the computer between the first data collection for analysis of consultations and the data collection for comparison with videorecorded data; doctors did not appear to change their recording habits for the prospective videorecorded validation study.

Proportionately more topics were captured in the manual records than were recorded on computer (Table 4). One exception was for the recording of diagnoses; these were more commonly recorded on the computer than in the manual notes or seen on the videorecordings (chi square, P<0.001). Some diagnoses were recorded on the computer as having been reviewed in the consultation but, under the rules for assessing the videorecordings, there was no overt discussion of that diagnosis. Another exception was for the recording of prescriptions; these were sometimes issued with no verbal interchange to indicate that this had occurred (chi square, P<0.001). Referral data were comparable in all three media.

Table 4. Topics occurring in 200 videorecorded consultations in one practice and percentage of those topics recorded as items in manual records or on computer.

Topic	% of occurrences of topic (item) recorded in records	
	Manual	Computer
Symptoms (n = 290)	53.8	5.5
Diagnosis (n = 161)	65.8	131.7
Prescription (n = 145)	83.4	101.4
Review arranged (n = 104)	35.6	5.8
Duration of symptoms (n = 91)	48.4	1.1
Site of symptoms (n = 72)	50.0	5.6
Preventive advice (n = 64)	37.5	7.8
Numerical finding (n = 63)	93.7	25.4
Previous test results (n = 41)	68.3	22.0
Investigation arranged (n = 35)	68.6	8.6
Referral (n = 15)	100	93.3
Preventive action (n = 8)	50.0	50.0
Treatment given (n = 8)	62.5	62.5

n = number of occurrences of topic in the videorecordings.

Discussion

If the data from general practice computer systems are to be of value to the practices and, in aggregated form, to the health service, they must be complete and accurate. In these four practices, acknowledged to be especially interested in and committed to data recording, the evidence is cautiously encouraging. In particular, practice D appeared to be using its computer for extensive consultation recording in addition to its manual records, and this pattern can be expected to become more common when the legal barriers to using only computer records have been overcome.

In all four practices the computer recording of 10 major diagnoses appeared to be sufficiently complete and accurate to be of value. No practice consistently, across all 10 diagnoses, recorded prevalences higher or lower than the other practices, suggesting that these four practices were either recording such morbidity accurately or with systematic and similar errors. It was gratifying to note consistency in prevalence rates for asthma, a disease with problems in definition. However, subjective diagnoses were less consistently recorded than more objective diagnoses and thus caution should be taken in interpreting data from practices about subjective areas.

From this study, it appears that two particular issues must be addressed before aggregated morbidity data can be relied upon. First, there is a need to use multiple search strategies, for example, searching for trabeculectomy in addition to glaucoma, which could be simplified by standardizing data entry. Secondly, there is a need for standardized disease inclusion and exclusion criteria, for example, to define the end of active asthma, which would lead to greater consistency between practices. However, in this study the use of a list of patients having had a prescription for nitrates in the previous two years to assess the extent of coronary heart disease would have missed those patients with coronary heart disease who were symptom free (for example, having had a myocardial infarction some years before or coronary artery surgery) and those on effective prophylaxis with drugs other than nitrates.

The recording of lifestyle data was poor and there is no agreement on how often patients should be asked about their lifestyles. Is a recording five years previously of 'non-smoker' satisfactory? It might be in a patient aged 70 years, but not in a patient aged 20 years. In time, lifestyle data should become sufficiently complete to be usable, since practices are increasingly recording these data for health promotion bandings. Socioeconomic data, such as occupation, social class and ethnicity, were notably absent. Census data can be used through small area statistics to make good this shortfall at practice level but not at an individual patient level.[13]

Comparison of the manual records with the computer records in the four practices suggested that for the main items of interest in data aggregation — diagnoses, prescriptions and referrals — the computer files were either more accurate than or were comparable to the manual records. However, it was clear that practices were selecting areas that they considered important to record on their computer systems; where their decisions did not coincide, wide variations in recording rates were found.

The videorecording of consultations in one practice showed that, as expected,[17] the act of videorecording did not appear to affect doctor behaviour. Many topics seen on the videorecording were not recorded in the manual records or on computer, but these were not topics of clinical importance.

The results of this study suggest that these four practices are recording to a sufficient level of completeness and accuracy for their computer files to be of value for assessing prevalence, prescribing and referral data, but not for reflecting the rich detail of the consultations as recorded in traditional manual records.

References

1. Taylor M, Ritchie L, Taylor R, et al. General practice computing in Scotland. BMJ 1990; **300:** 170-172.
2. Jick H, Jick S, Derby L. Validation of information on general practitioner based computerised data resource in the United Kingdom. BMJ 1991; **302:** 766-768.
3. Johnson N, Mant D, Jones L, Randall T. Use of computerised general practice data for population surveillance: comparative study of influenza data. BMJ 1991; **302:** 763-765.
4. Pringle M, Hobbs R. Large computer databases in general practice [editorial]. BMJ 1991; **302:** 741-742.
5. Gilliland A, Mills K, Steele K. General practice records on computer — handle with care. Fam Pract 1992; **9:** 441-450.
6. Laporte R. Global public health and the information superhighway [editorial]. BMJ 1994; **308:** 1651-1652.
7. Secretary of State for Health. The health of the nation: a strategy for health in England (Cm 1986). London: HMSO, 1992.
8. Chisholm J. The Read clinical classification [editorial]. BMJ 1990; **300:** 1092.
9. Allen K, Markwell D. MIQUEST project report. Morpeth: Northumberland Health, 1994.
10. Walton R, Randall T. Communication in the year 2000 [editorial]. Br J Gen Pract 1994; **44:** 434-435.
11. Nazareth I, King M, Haines A, Rangel L. Accuracy of diagnosis of psychosis on a general practice computer system. BMJ 1993; **307:** 32-34.
12. Yeoh C, Davies H. Clinical coding: completeness and accuracy when doctors take it on [editorial]. BMJ 1993; **306:** 972.
13. Ward P, Morton-Jones A, Pringle M, Chilvers C. Generating social class data in primary care. Public Health 1994; **108:** 279-287.
14. Pringle M. Greeks bearing gifts [editorial]. BMJ 1987; **295:** 738-739.
15. Royal College of General Practitioners, Office of Population Censuses and Surveys, and Department of Health. Morbidity statistics from general practice. Fourth national study 1991-92. London: HMSO, 1995.
16. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measurement 1960; **20:** 37-46.
17. Pringle M, Stewart-Evans C. Does awareness of being video recorded affect doctors' consultation behaviour? Br J Gen Pract 1990; **40:** 455-458.

Address for correspondence

Professor M Pringle, Department of General Practice, Nottingham University Medical School, Queens Medical Centre, Nottingham NG2 7UH.

Food for thought...

'The incidence of hospital admissions for self-inflicted injury/poisoning was closely related to the socioeconomic conditions prevailing in the area where the practice was located. A typical four-doctor practice in an affluent area with less than 5% male unemployment might have one hospital admission for self-inflicted injury/poisoning every two months. In contrast, a similarly sized practice located in an area with male unemployment at over 15% might expect to have one such hospital admission every two weeks.'

Smith T. Differences between general practices in hospital admission rates for self-inflicted injury and self-poisoning: influence of socioeconomic factors. September *Journal*, p.458.