The quality of record keeping in primary care: a comparison of computerised, paper and hybrid systems

William T Hamilton, Alison P Round, Deborah Sharp and Tim J Peters

SUMMARY

Background: Computerised record keeping in primary care is increasing. However, no study has systematically examined the completeness of computer records in practices using different forms of record keeping.

Aim: To compare computer-only record keeping to paper-only and hybrid systems, by measuring the number of consultations and symptoms recorded within individual consultations.

Design of study: Retrospective cohort study.

Setting: Eighteen general practices in the Exeter Primary Care Trust.

Method: This study was part of a retrospective case control study of cancer patients aged over 40 years. All recorded consultations for a 2-year period were identified and coded for 1,506 patients. Records were classified as paper, computer, or hybrid, depending on which medium stored the clinical information from consultations.

Results: More consultations were recorded in hybrid systems (median in 2 years = 11, interquartile range [IQR] = 6–18) than computer systems (median in 2 years = 9, IQR = 4–16,5) or paper systems (median in 2 years = 8, IQR = 5–14,): P<0.001. In a Poisson regression analysis, which included age, sex, and history of cancer diagnosis, the rates of consultations recorded in paper and computer systems were 16% and 11% lower, respectively, than in hybrid systems. Fewer telephone consultations were recorded in paper systems, and fewer home visits in computer systems. Fewer symptoms were recorded in individual consultations on computer systems. Recording of absent symptoms and severity of symptoms was highest in paper systems.

Conclusion: Hybrid systems of primary care record keeping document higher numbers of consultations than computer-only or paper-only systems. The quality of individual consultation recording is highest in paper-only systems. This has medico-legal implications and may impact upon continuity of care.

Keywords: medical records systems, computerised; cohort studies; regression analysis.

W T Hamilton, FRCP, FRCCP, research fellow; D J Sharp, PhD, FRCCP, professor of primary health care; T J Peters, PhD, FFPHM, FRCCP, professor of primary health services research, Division of Primary Health Care, University of Bristol. A P Round, FFPHM, MRCP, MRCGP, director of public health, East Devon Primary Care Trust, Exeter.

Address for correspondence:
Dr William Hamilton, Division of Primary Health Care, University of Bristol, 1 Gaff House, Bristol BS6 6LL.
E-mail: w.t.hamilton@btopenworld.com


Introduction

Most general practices now use computers for their record keeping.1,2 This has many advantages, principally because computerised records are easier for retrieval of information. Thus, maintaining age–sex and disease registers, organising recall systems, and performing audits are all simplified.3 However, for information to be extracted from a computerised system, it first has to be entered. The completeness of information is crucial to the value of the system. High quality record keeping is useful for continuity of care, and may be essential for medicolegal purposes.

Very little research has examined the quality of data entry in computerised records.4–7 Three United Kingdom (UK) studies have compared computerised to paper primary care records. The first study examined the completeness of data in four practices that were selected for their high computer usage, but which also maintained paper records.6 In the paper records, recording of symptoms and their duration was much higher, and more topics were recorded. A second study, again in a practice using paper and computer records in parallel, demonstrated a lower level of recording of investigation results in the computerised system.8 One older study showed an under-reporting of influenza in practices using computerised systems when compared to illness records collected manually.9

No study has systematically examined the completeness of data recording between practices using different systems of record keeping. We therefore designed an observational study to compare paper-only systems, computer-only systems and systems incorporating both paper and computerised records. For simplicity these are referred to as ‘paper’, ‘computer’, and ‘hybrid’ systems in the rest of this paper.

Method

This study was part of a retrospective case control study examining general practitioner records for possible symptoms that may predict cancer. Of the 21 practices in the Exeter Primary Care Trust, 18 participated in this study. All patients over the age of 40 years with either prostate, lung or colorectal cancer were identified from the local hospital cancer register during the period from January 1998 to September 2001. For each patient with cancer, five age, sex, and general practice-matched controls were identified from the practices’ computerised age–sex registers.

The entire primary care records were copied and anonymised for the 2-year period before the diagnosis of cancer in both cases and their matched controls. Records of primary care consultations were classified into three groups: computer, paper, or hybrid. Classification was based on
which medium stored the clinical entry for the consultation. If, for example, the consultation details were entered on paper but a prescription without clinical data was generated through the computer, this was still classified as paper. If, however, the computer entry in such cases also included a diagnosis or symptoms, then this was classified as hybrid. The hybrid group therefore largely comprised those with parallel recording of clinical information in both media. In a small number of patients the recording system had changed during the 2 years of study. These transitional patients were also classified as hybrid, as there was usually a period of parallel recording during the change. One practice used A4 paper records, with the remainder using Lloyd George cards. The computer systems in use were: EMIS (9), Exeter (3), VAMP (3), Meditel (2), Microtest, System 2000, and Surgery Manager. Two practices had changed systems during the study.

All consultations were entered onto a specially designed database. Consultations were classified by their site, i.e. surgery appointment, home visit, telephone, prescription only, certificate only, or out-of-hours. Only one consultation was counted per day, and this was reinforced by designing the database to block any inadvertent duplicate entries. Clinical information for each consultation was coded using an adapted form of the International classification of primary care-2 (ICPC-2) (Box 1).

Trained research assistants coded both symptoms and diagnoses if both were recorded. All coded a mixture of records from the three recording systems. Both doctor and nurse consultations were coded, as paper notes were generally unsigned, and computerised records were marked by initials, but did not indicate the professional background of the person doing the recording. The outcome measures were: the number and setting of consultations recorded in the three systems; the number of symptom codes recorded per consultation; and the number of symptom codes recorded with a negation, duration or severity statement. The full study was approved by the Exeter Local Research Ethics Committee.

Data analysis
Descriptive statistics and 95% confidence intervals (CIs) for percentages were calculated. Consultation numbers and the number of codes within consultations were not normally distributed, so differences between recording systems were analysed by medians and Kruskal–Wallis tests, including pairwise comparisons. Poisson regression was performed to examine the effect of future cancer diagnosis, sex and age on recorded consultation rates. The types of consultations, and the quantification of symptom codes, were analysed by Pearson $\chi^2$-tests. STATA 7 software was used for all analyses.

Results
Consultation recording
At the time of study closure, 1396 patients had been fully coded, and their details and consultation numbers are shown in Table 1. The overall difference in the number of recorded consultations was highly significant across the three recording systems. Pairwise comparisons, adjusted for multiple testing, were also significant for the hybrid versus the paper systems ($P<0.001$) and for the hybrid versus the computer systems ($P = 0.02$), but not for the computer versus the paper systems ($P = 0.30$). A Poisson regression analysis was performed, which included age, future cancer diagnosis, sex, and record type (Table 2). This shows that differences between the record types remained highly significant after controlling for the other factors. The rates of recorded consultations in paper and computer systems were 16% and 11% lower than for hybrid systems, respectively. The types of consultations — telephone, surgery, or home visit — are shown in Table 3.

Symptom recording
This was restricted to surgery consultations that recorded at least one symptom code. Thus, procedural consultations such as blood tests and immunisations were omitted. A further 144 consultations with an indecipherable entry were also omitted (2 computer, 122 paper, 20 hybrid).

Details of symptom recording in these consultations are shown in Table 4. More symptom codes per consultation were recorded in paper and hybrid systems. Negative symptom recording, and recording of the severity, was highest in...
Table 1. Number and characteristics of patients whose records were studied.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Paper System</th>
<th>Computer System</th>
<th>Hybrid System</th>
<th>Total</th>
<th>P-value (type of test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>713</td>
<td>321</td>
<td>362</td>
<td>1396</td>
<td></td>
</tr>
<tr>
<td>Median age in years (IQR)</td>
<td>69.4 (61.4–76.3)</td>
<td>73.1 (67.6–77.5)</td>
<td>72.5 (65.2–76.9)</td>
<td>71.3 (63.6–76.9)</td>
<td>&lt;0.001 (Kruskal–Wallis)</td>
</tr>
<tr>
<td>Number (%) female</td>
<td>254 (35.7)</td>
<td>117 (36.6)</td>
<td>131 (36.2)</td>
<td>502 (36.0)</td>
<td>0.96 (Pearson χ², 2 df)</td>
</tr>
<tr>
<td>Future cancer cases (%)</td>
<td>104 (14.6)</td>
<td>59 (18.4)</td>
<td>79 (21.8)</td>
<td>242 (17.4)</td>
<td>0.01 (Pearson χ², 2 df)</td>
</tr>
<tr>
<td>Total number of consultations recorded</td>
<td>7361</td>
<td>3705</td>
<td>4692</td>
<td>15758</td>
<td></td>
</tr>
<tr>
<td>Median (IQR) number of consultations per patient</td>
<td>9 (5–14)</td>
<td>10 (4–16)</td>
<td>11 (6–18)</td>
<td>9 (5–16)</td>
<td>&lt;0.001 (Kruskal–Wallis)</td>
</tr>
</tbody>
</table>

df = degrees of freedom. IQR = interquartile range.

Table 2. Poisson regression analysis of characteristics affecting consultation numbers.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Adjusted ratio of mean number of consultations</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex</td>
<td>0.79</td>
<td>0.76 to 0.81</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age</td>
<td>1.01a</td>
<td>1.009 to 1.012</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Having a future cancer diagnosis</td>
<td>1.27</td>
<td>1.32 to 1.43</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Record system</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hybrid</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper records only</td>
<td>0.84</td>
<td>0.81 to 0.87</td>
<td></td>
</tr>
<tr>
<td>Computer records only</td>
<td>0.89</td>
<td>0.86 to 0.93</td>
<td></td>
</tr>
</tbody>
</table>

*For a 1-year increase in age.

Table 3. Setting of consultations recorded in the three systems.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Paper (n = 711)</th>
<th>Computer (n = 320)</th>
<th>Hybrid (n = 361)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients who had a telephone consultation (% of cohort)</td>
<td>111 (16)</td>
<td>87 (27)</td>
<td>103 (29)</td>
</tr>
<tr>
<td>Patients who did not have a telephone consultation but did have a home visit (% of cohort)</td>
<td>78 (11)</td>
<td>19 (6)</td>
<td>30 (8)</td>
</tr>
<tr>
<td>Patients who did not have a telephone or home visit but did have a surgery consultation (% of cohort)</td>
<td>522 (73)</td>
<td>214 (67)</td>
<td>228 (63)</td>
</tr>
</tbody>
</table>

*Four patients had not had a consultation of one of these types. The differences between the groups are significant: P<0.001, Pearson χ², 4 df.

Table 4. Details of symptom recording in consultations containing at least one symptom code.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Paper</th>
<th>Computer</th>
<th>Hybrid</th>
<th>Total</th>
<th>P-value (type of test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>638</td>
<td>278</td>
<td>331</td>
<td>1247</td>
<td></td>
</tr>
<tr>
<td>Number of consultations</td>
<td>2914</td>
<td>1471</td>
<td>1860</td>
<td>6245</td>
<td></td>
</tr>
<tr>
<td>Number of individual symptom codes</td>
<td>5222</td>
<td>2444</td>
<td>3356</td>
<td>11022</td>
<td></td>
</tr>
<tr>
<td>Median (mean) number of codes per consultation</td>
<td>1 (1.79)</td>
<td>1 (1.66)</td>
<td>1 (1.80)</td>
<td>1 (1.76)</td>
<td>0.002 (Kruskal–Wallis)</td>
</tr>
<tr>
<td>Number of negative codes (%)</td>
<td>656</td>
<td>225</td>
<td>344</td>
<td>1225</td>
<td>&lt;0.001 (Pearson χ², 2 df)</td>
</tr>
<tr>
<td>Number of codes with severity recorded (%)</td>
<td>770</td>
<td>290</td>
<td>327</td>
<td>1387</td>
<td>&lt;0.001 (Pearson χ², 2 df)</td>
</tr>
<tr>
<td>Number of codes with duration recorded (%)</td>
<td>465</td>
<td>227</td>
<td>370</td>
<td>1062</td>
<td>0.004 (Pearson χ², 2 df)</td>
</tr>
</tbody>
</table>

df = degrees of freedom.
the paper systems. Recording of the duration of a symptom was highest in the hybrid systems. Three symptoms potentially associated with cancer were studied further: D16 (rectal bleeding); R24 (haemoptysis), and U06 (haematuria). These were recorded 266 times, 72 of which were in the negative. Recording of the absence of these symptoms was highest in the paper systems: 41/127 (32%) codes, compared to 12/73 (16%) in the computer systems, or 19/66 (29%) in the hybrid systems ($P = 0.05$, Pearson $\chi^2$-test).

**Secondary analyses**

The lower number of recorded consultations in paper and computer systems could have two main explanations: first, the number of actual consultations is the same but fewer are recorded, or, second, that practices with different record keeping systems have different access policies resulting in fewer actual consultations. To examine the second possibility, analyses were repeated in practices that had changed their record keeping method during the study. It is unlikely that such practices would have also changed their access policies at the same time. Over the 7-year period of the study, six practices (414 patients) had paper systems throughout and four (246 patients) had computer systems throughout. Eight practices had made a transition: four from paper to hybrid systems, two from hybrid to computer systems, and two from paper, through hybrid, to computer systems. These transitional practices provided 734 patients. In these transitional practices, the median number (interquartile range [IQR]) of consultations in computer systems was 10 (4–17), in paper systems it was 8 (4–13), and in hybrid systems it was 11 (6.5–18), with $P<0.001$, using the Kruskal–Wallis test.

**Discussion**

**Main findings**

This study shows significant differences in general practice record keeping, depending on whether the practice uses paper records, computerised records, or a hybrid system. With hybrid systems as a reference standard for this study, computerised systems recorded fewer home visits, and paper systems fewer telephone consultations. Within the consultation itself, more symptoms are recorded and more symptoms are quantified when paper systems are used. This accords with the two previous UK studies comparing computerised to manual records. However, by studying hybrid practices, these papers may simply have demonstrated that clinicians keep better records when both options are kept in parallel.

Some of our findings are simple to explain. Telephone consultations are easier to record on a computer, as the computer will usually be on the same desk as the telephone — this will rarely be the case for the paper records. It may be that some telephone consultations are recorded in a separate ‘call-back’ book in paper practices. These would have been missed in this study. Conversely, few practices use portable computers, so a doctor performing a home visit will have to remember to record it on return to surgery.

Within individual consultations, paper systems record more symptoms for each consultation than computer systems. They are also better at recording the severity of symptoms, and in specifically documenting absent symptoms. One reason could be that clinicians find it easier to write on paper than with a keyboard, and so record more. Some computer systems used in the study had limited free text (VAMP in particular), and others allowed a positive symptom to be recorded in the coding system, but not a negative symptom. Some of the written symbols were commonly accepted medical shorthand, such as ‘+ +’ (meaning ‘moderate’) or ‘P’ (meaning ‘no’ or ‘not’). These notations, although possible on a keyboard, may not be easily used. All these points would encourage minimalist record keeping with computer systems. As expected, paper systems had more indecipherable entries, with about 2% of paper records being illegible. This is less than popular folklore would suggest, but still of importance.

**Strengths and limitations**

This is a large study, examining 2792 patient-years of records and 15 758 consultations. The principal assumption is that patients from different practices will have the same consultation rates and make the same number of telephone and home consultations. There are large overlaps in the geographical area served by Exeter practices, so it is unlikely that socioeconomic factors have an important effect on the results. Practices which have chosen to become paperless may have different access policies that encourage telephone consultations and discourage home visits. We have no evidence of this. Furthermore, this study does not address general practitioner characteristics. Doctors who have chosen to retain an element of paper recording may differ from those who have moved to a paperless practice. There is no published evidence to suggest that practices with computerised recording offer shorter consultations (and thus have less time to take good notes), although this is possible. The secondary analysis, only examining practices that changed their recording method during the study, showed very similar results to the study as a whole, which adds support to our interpretation that it is the recording of consultations that differs between the cohorts, not the actual number of consultations.

**Implications**

These results have medicolegal implications, as well as implications for continuity of care, for researchers and for compilers of information about primary care usage. The medicolegal implications are clear. If a consultation is unrecorded, a doctor will have lost a major piece of evidence, should that consultation lead to a complaint. This applies as much, or possibly more, to home visits and telephone consultations. Continuity of care will suffer, too; without a record of the previous consultation important follow-up items may be missed. This problem is magnified if a different doctor takes the second consultation. Researchers who base their studies on recorded consultations will need to consider our findings when analysing their studies. Similarly, consultation rates derived from recorded consultations will be underestimated, particularly if coming from paper records.

The transition to computerised record keeping in primary care is set to continue, and the advantages are great.
Indeed, it is probable that before computerised record keeping very few telephone consultations were recorded, so this study shows one of the benefits of such records. However, if the move towards being paper-free is accompanied by a reduction in the quality of record keeping of individual consultations, the disadvantages may also be considerable. This could be prevented in part if computer systems were developed to include items such as severity or duration. Until this happens, primary care clinicians will need to be careful to ensure that their records retain high quality as well as high accessibility.

References


Acknowledgements

We wish to thank all Exeter practices for their help, and S Stapley, J Barrett, B Laws, H Massie, C Stabb and A Waterlow for their enthusiastic work. The main study was funded by the Department of Health, and Dr William Hamilton was funded through his research practice in Barnfield Hill, Exeter, and RCGP/BUPA and National Health Service Fellowships.

Commentary

This paper uses Poisson regression to compare consultation rates in practices using computer-only record keeping to those using paper-only or hybrid systems. The mathematical distribution underlying this method is called the Poisson distribution; it was named after the French mathematician Siméon-Denis Poisson (1781–1840). However, an understanding of the Poisson distribution is unnecessary for users of the method.

Poisson regression is used to estimate event rates and, more importantly, to estimate rate ratios comparing different groups. For example, we might estimate the rate of breast cancer among women aged 50–60 years registered in a general practice, or we might estimate rate ratios comparing rates of asthma attacks in children living in households that are adjacent and not adjacent to a main road. To use Poisson regression, we require two pieces of information for each individual or group in our dataset: the number of times that the event occurred (this may be zero), and the total follow-up time. The rate of the event is estimated as:

\[
\text{rate} = \frac{\text{total number of events}}{\text{total follow-up time}}
\]

Epidemiologists have known for many years how to estimate rates and rate ratios, with associated confidence intervals and \( P \)-values. While Poisson regression provides a simple and convenient way to do this, its main advantage is that it allows us to control for the possible confounding effects of other factors. For example, if children living adjacent to main roads tend to be from poorer households than other children, and household wealth is associated with asthma rates, we would want to take account of this before asserting that living adjacent to a main road is associated with increased rates of asthma attacks. Because of their convenience and flexibility, Poisson regression analyses are replacing standardisation of rates as a means of controlling for the confounding effects of factors such as age and sex.

A note of caution is required. Poisson regression analyses are not valid for clustered data, particularly when events may occur more than once in an individual, or when events are grouped. For example, a single child may experience multiple asthma attacks over a 1-year period. A Poisson regression analysis that fails to allow for clustering will tend to produce confidence intervals that are too narrow and \( P \)-values that are too small: this may lead researchers to believe, wrongly, that they have found evidence of an association. Fortunately, an extension to Poisson regression, called negative-binomial regression, is available in a number of statistical packages. This is based on a generalisation of the Poisson distribution that incorporates the clustering effects.

Clayton and Hills provide excellent explanations of the mathematical basis of Poisson regression.\(^1\) Kirkwood and Sterne give practical examples of the use of the method.\(^2\)

Jonathan A C Sterne
Reader in Medical Statistics and Epidemiology, Department of Social Medicine, University of Bristol.

References
