Does Feedback Improve the Quality of Computerized Medical Records in Primary Care?

SIMON DE LUSIGNAN, BSc, MBBS, PETER N. STEPHENS, BA, MA, NAEEMA ADAL, BSc, AZEEM MAJEED, MB BCH, MD

Abstract Objective: The MediPlus database collects anonymized information from general-practice computer systems in the United Kingdom, for research purposes. Data quality markers are collated and fed back to the participating general practitioners. The authors examined whether this feedback had a significant effect on data quality.

Methods: The data quality markers used since 1992 were examined. The authors determined whether the feedback of “useful” data quality markers led to a statistically significant improvement in these markers. Environmental influences on data quality from outside the scheme were controlled for by examination of the data quality scores of new entrants.

Results: Three quality markers improved significantly over the period of the study. These were the use of highly specific “lower-level” Read Codes ($p=0.004$) and the linkage of repeat prescriptions ($p=0.03$) and acute prescriptions ($p=0.04$) to diagnosis. Clinicians who fall below the target level for linkage of repeat prescriptions to diagnosis receive more detailed feedback; the effect of this was also statistically significant ($p<0.01$).

Conclusions: The feedback of four of the ten markers had a significant effect on data quality. The effect of more detailed feedback appears to have had a greater effect. The lessons learned from this approach may help improve the quality of electronic medical records in the United Kingdom and elsewhere.

Until now, only clinicians who volunteered have been part of data quality schemes and received feedback on the quality of their coding. There is no clear evidence about whether such feedback can improve data quality. For example, the effectiveness of PCO (primary care organization)-wide feedback on data quality has yet to be shown by the Primary Care Organization Data Quality (PCDQ) program.\textsuperscript{5}

We examined the feedback of data quality markers within the MediPlus database to see whether this led to a more rapid improvement in data quality than that generally occurring in primary care. We hoped that the experience gained from data quality feedback over an 8-yr period could be applied more generally to raising the standards of computerized medical records in primary care.

**Methods**

**Data Source: The MediPlus Database**

The MediPlus database was established in 1992; it contains information on almost two million patients and more than 53 million prescriptions.\textsuperscript{16} The database is based on information drawn from more than 500 representative general practitioners across the United Kingdom using the Torex-Meditel System 5 computer package.\textsuperscript{17} This computer system allows the linkage of diagnosis or problem title to the acute or repeat (long-term) prescriptions issued to patients. This makes clearer the diagnoses for which prescriptions are being issued. This is particularly useful among groups such as the elderly, who often suffer from several chronic diseases.

Data quality markers are used to ensure that only doctors supplying data that reaches specified quality standards are included in the database used by researchers. In total, ten data quality scores are used. These are calculated at individual doctor level and fed back to the participating practices quarterly. Newsletters are also sent every six months, addressing issues around coding highlighted by a panel of expert general practitioners. Doctors are given a small incentive (about £400 per doctor per year) to reach the target levels across the ten quality scores used (Table 1).

**Literature Review**

We carried out a literature review to establish whether there was a consensus on what data quality parameters should be fed back. PubMed (National Library of Medicine) was searched using “Data Quality” and “General Practice” as search terms. This identified 848 abstracts, each of which was examined to identify articles relating to either the membership of a data quality scheme or the effectiveness of feedback of clinical markers. There were many descrip-
Table 1

Data Quality Markers Used by the MediPlus Database

<table>
<thead>
<tr>
<th>Quality Marker</th>
<th>Reason for Its Inclusion</th>
<th>Weakness as a Marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Percentage of registered patients for whom there has been a change in the record over the previous 12 mo</td>
<td>An indication that that system is being used routinely.</td>
<td>Less useful since Health Authority registers and practice data bases are linked (GP-Links Project)</td>
</tr>
<tr>
<td>2. Percentage of patients with year of birth and sex recorded</td>
<td>Ensures that researchers can analyze disease by age and sex of patient</td>
<td>Less useful since Health Authority registers and practice data bases are linked (GP-Links Project)</td>
</tr>
<tr>
<td>3. Percentage of problems or diagnoses with Read Code of level 3 or lower</td>
<td>Lower-order codes represent more specific diagnoses. Some high-order codes contain negatives within the lower orders.</td>
<td>May contradict primary care group choices about coding data.</td>
</tr>
<tr>
<td>4. Percentage of notes linked to problem or diagnosis</td>
<td>Linkage of notes increasingly important for analysis of test results</td>
<td>Electronic transmission of test results cannot be linked to a problem but only pasted into the patient’s notes.</td>
</tr>
<tr>
<td>5. Percentage of notes in which Read Code is level 3 or lower</td>
<td>As marker 3</td>
<td>May contradict Primary Care Group choices about coding data.</td>
</tr>
<tr>
<td>6. Number of prescriptions issued per week per 1000 registered patients</td>
<td>This is a crude measure of how much prescribing is not being computerized. Also looking for abnormalities in trends over time that would allow detection of missing data</td>
<td>Duration of repeat prescription interval can seriously affect this marker, e.g., a practice that begins to issue repeat prescription for 3-mo intervals, as opposed to 1- or 6-mo intervals, would see radical changes.</td>
</tr>
<tr>
<td>7. Complete dose and regimen details related to dose-effect or ADR</td>
<td>Important for prescribing analyses</td>
<td>Once acute prescribing is computerized, it does not differentiate between practices.</td>
</tr>
<tr>
<td>8. Proportion of acute prescriptions issued linked to a problem title or diagnosis</td>
<td>A key function of the database is to show what the prescribing behavior of general practitioners is.</td>
<td>Office automation may drive this process. Once these are being issued, specifically targeting home visits may be more important.</td>
</tr>
<tr>
<td>9. Proportion of repeat prescriptions linked to a problem title or diagnosis</td>
<td>A key function of the database is to show what the prescribing behavior of general practitioners is.</td>
<td>When practices newly join the scheme, doing this linkage de novo may lead to inaccuracies.</td>
</tr>
<tr>
<td>10. Ratio of acute prescriptions issued to chronic prescriptions</td>
<td>Checks for consistent usage</td>
<td>See comments for 5, 6, and 7 above.</td>
</tr>
</tbody>
</table>

Improvements in the quality marker scores may have been due to various NHS initiatives, such as Collection of Health Data from General Practice. If general practitioners were improving “naturally,”
this would be reflected in an increase in the starting scores of general practitioners who joined the scheme over time. General practitioners were therefore grouped according to year of joining, their starting scores on each marker extracted, and regression analysis on the means for each group used to see whether their starting scores improved over time on any marker.

Excluding the Effect of Differential General Practitioner Drop Out

Results may be biased by a greater proportion of the poorly performing doctors dropping out during the early years of the scheme. To check this, doctors were first grouped according to the length of time they had spent in the scheme, regardless of start date. For example, two general practitioners who started in 1992 and in 1994 but who remained in the scheme for three years would be placed in the same group. The difference in each general practitioner’s first and last scores was calculated and from these the mean scores were found for each group and marker. Regression analysis was again used to determine whether time in scheme affected data quality.

Effect of Specific Feedback to Those with a Below-average Score for Linkage of Diagnosis and Prescription

One particular form of feedback was also investigated, an initiative to improve the linkage of diagnosis to prescription. Its effectiveness was assessed by comparing the mean score for all the general practitioners in the second quarter of 1999 who received the report, with their mean score in the first quarter of 2000.

Results

Three Markers Show Improvement with Time

The quality markers showing a significant improvement with time at the 5 percent level were:

- Percentage of acute prescriptions linked to a diagnosis
- Percentage of repeat prescriptions linked to a diagnosis
- Percentage of problems defined by a Read Code of level 3 or lower.

The mean starting scores and the results of the regression analysis are shown in Tables 2 and 3.

External Environmental Factors Do Not Explain Improvement

None of the markers that improved over time showed any evidence that external factors had played a part. Table 4 shows the results of the regression analysis on the general practitioners’ starting scores. The only marker showing any significant improvement in starting score with time was the number of prescriptions issued per 1,000 registered patients ($p<0.05$).
Differential General Practitioner Drop Out Cannot Explain Improvement

It was possible to do this analysis for two data quality markers—the percentage of repeat prescriptions and the percentage of acute prescriptions linked to a diagnosis. The repeat prescription linkage showed significant improvement with time in the scheme, regardless of the general practitioner start date ($p < 0.05$).

Acute prescription linkage did not show any such trend. However, whether time spent in the scheme was short or long, acute prescription linkage improved (range, 2.5–15.1 percent). These two analyses suggest that the improvements seen in coding quality were not due to differences in the rate at which poorly performing general practitioners dropped out of the scheme.

**Significant Effect of Specific Feedback for Linkage of Diagnosis and Prescription:**

The feedback of the detailed repeat prescribing reports had a significant effect on the percentage of repeat prescriptions linked to diagnosis. This increased from 64.6 percent in the second quarter of 1999 to 79.5 percent in the first quarter of 2000 (difference, 14.9 percent; 95% confidence interval, 3.7–26.1 percent; $p = 0.005$).

**Discussion**

The main finding from this study is that the feedback of four of the quality markers improved data quality. All these markers were fed back over a long period; one marker was also fed back over a shorter period, with the specific aim of increasing the linkage of diagnosis or problem to repeat prescriptions. The scheme members were also offered a small financial incentive, but this was dependent on meeting target scores for all markers, not specific ones. To receive this payment, general practitioners would have needed to focus on those markers for which they performed least well.

The results of this study are mixed. Feedback of half the markers achieved significant improvement, while feedback of others did not. Feedback of this nature is not, therefore, in itself an effective mechanism, but it
may represent a low-cost tool that can be used alongside other tools. The explanations for why the short-term feedback was so successful also needs to be explored further.

The findings from this study have potentially important implications for electronic patient records. First, they can inform those seeking mechanisms to improve data quality about the effects of a long period of feedback to a large group of practices. Second, those feeding back data quality indexes may wish to critically examine whether the feedback had any effect on data quality. Third, they indicate the importance of further research to describe the context in which feedback may contribute toward improvement in data quality.

Some potential confounding factors need to be considered. The members of the MediPlus database all use the Meditel computer system and are volunteers. They may have made considerable efforts to raise their data quality standards before joining the scheme. Some general practitioners had data quality markers that were already at levels over 90 percent before they joined the scheme. For markers with such high scores, it would be difficult to show a statistically significant improvement in score over time. Finally, some markers have been overtaken in their usefulness by advances in technology. Automated registration links with the Health Authority (GP-Links Project) has almost eliminated the number of patients without full demographic details. Similarly, patients who die or move away are now more likely to be automatically removed from a practitioner’s list. In the past, unremoved patients may have artificially increased the list size, thereby increasing the denominator population used to calculate the data quality scores. Technical solutions like GP-Links have clearly had a major influence on some aspects of general practitioner data.

Further research is needed to ascertain what data quality markers should be fed back and by whom. From the literature review, individual feedback on a narrow clinical focus seems to offer the best approach to feedback of data quality markers. However, this was not the mechanism used for the most successful data quality marker fed back within the MediPlus database. Research is also needed to ascertain whether the personal feedback, the token financial rewards, or some other factor was responsible for this change. The role of practice staff may also need to be carefully examined, as primary care support staff are responsible to varying degrees for issuing repeat prescriptions.

**Conclusions**

We found that four data quality markers, all relating to the linkage of diagnosis to prescription and the use of more specific Read Codes, improved at a significantly higher rate in MediPlus practices. The personalized feedback to those general practitioners with below-average scores and the token financial incentives may have been important motivating factors and should be tested elsewhere. However, the role of practice support staff and the improvements made to the accuracy of the denominator through the GP-links project show that factors other than coding by clinical staff may have profound effects on data quality. If general practice computer records are to become the cornerstone of the electronic patient and health records promised by the NHS information strategy, research is urgently needed to define how feedback on data quality should be given.

Simon de Lusignan heads a Primary Care Informatics Group within the Department of General Practice and Primary Care at St. George’s Hospital Medical School. Peter Stevens and Naema Adal are employees of IMS, which runs the MediPlus database. Azeem Majeed holds a Primary Care Scientist Award and is funded by the NHS Research and Development Directorate. Neither Simon de Lusignan nor Azeem Majeed received any payment for this study, and they have no financial interest in IMS.

**References**

9. Thiru K, de Lusignan S, Hague N. Have the completeness and

http://www.nhsia.nhs.uk/terms/pages/readcodes_intro.asp?om=m1
33. CHDGP (Collection of Health Data from General Practice), Nottingham, UK. Available at: http://www.nottingham.ac.uk/chdgp/. Accessed Jun 13, 2002.