

Primary care



This is an abridged version; the full version is on bmj.com

Fisher Medical Centre Research Unit, Skipton, North Yorkshire BD23 1EU

Krish Thiru
research fellow
Alan Hassey
general practitioner

Tayside Centre for General Practice, University of Dundee, Dundee DD2 4AD

Frank Sullivan
general practitioner

Correspondence to: K Thiru
krish.thiru@st-marys.nhs.uk

BMJ 2003;326:1070-2

Systematic review of scope and quality of electronic patient record data in primary care

Krish Thiru, Alan Hassey, Frank Sullivan

Abstract

Objective To systematically review measures of data quality in electronic patient records (EPRs) in primary care.

Design Systematic review of English language publications, 1980-2001.

Data sources Bibliographic searches of medical databases, specialist medical informatics databases, conference proceedings, and institutional contacts.

Study selection Studies selected according to a predefined framework for categorising review papers.

Data extraction Reference standards and measurements used to judge quality.

Results Bibliographic searches identified 4589 publications. After primary exclusions 174 articles were classified, 52 of which met the inclusion criteria for review. Selected studies were primarily descriptive surveys. Variability in methods prevented meta-analysis of results. Forty eight publications were concerned with diagnostic data, 37 studies measured data quality, and 15 scoped EPR quality. Reliability of data was assessed with rate comparison. Measures of sensitivity were highly dependent on the element of EPR data being investigated, while the positive predictive value was consistently high, indicating good validity. Prescribing data were generally of better quality than diagnostic or lifestyle data.

Conclusion The lack of standardised methods for assessment of quality of data in electronic patient records makes it difficult to compare results between studies. Studies should present data quality measures with clear numerators, denominators, and confidence intervals. Ambiguous terms such as "accuracy" should be avoided unless precisely defined.

Introduction

Accountability in the NHS is crucially dependent on the availability of high quality clinical information. This relies on the data collected.¹ A clear message emerging from government policy initiatives is the need for high quality data on health collected through electronic patient record (EPR) systems. The assessment of quality and improvement of primary care datasets has been repeatedly emphasised.² However, the criteria against which quality should be judged remain unclear.

We identified one review of mainly secondary care studies that described system and organisational

factors that affect quality of the data in EPR.³ We carried out a similar review but in primary care.

Methods

We searched all major bibliographic databases and several specialist datasets during the last quarter of 2001 (see bmj.com for databases and sources and web table A for search criteria).

We established a framework for categorising and selecting review papers (box). Eligible papers had to satisfy at least one aspect (numbered) of each category (A-C) within the box.

Framework for assessing eligibility of publications for review

All three categories (A-C) needed to be satisfied for a paper to be selected

A Reference standard

A modification of the "distance from patient" concept, which classified the reference standard used to judge quality³

- (1) Studies that used objective "close to patient" standards by using techniques such as video recording or direct examination
- (2) Studies that used interviews or questionnaire surveys of patient, next of kin, or their immediate carers as reference standard
- (3) Studies that used routine consultation data (databases, EPRs, paper records, discharge letter, etc) as standard reference
- (4) Studies that used national statistics or equivalent survey results as their reference standard

B Study objectives

- (1) Studies that measured change in EPR data quality or those that measured EPR data quality were classified as measuring data quality
- (2) Studies that used EPRs and commented on their quality were classified as scoping data quality

C Data types

Publications that investigate:

- (1) Diagnostic or symptom state of the patient
- (2) Patient management data—for example, health promotion, drug treatment, referrals, tests
- (3) Wider aspects of patient and practice management—for example, family history, ethnicity, socioeconomic status, immunisation, hospital episodes, consultation rates

Results

We identified 4589 abstracts and categorised 174 documents after primary exclusions. Of these, we included 47 journal publications, four reports, and one thesis from 1980-2001. Thirty seven studies measured data quality, and 15 used electronic patient records and commented on quality in the presence of a reference standard (scoping). These were analysed separately (table). Forty eight studies assessed diagnostic data, 20 assessed management information, and 13 examined wider aspects of routine data.

Measuring data quality

Thirty seven studies measured data quality, and 31 were from the United Kingdom. A similar proportion had been published since 1995. Table B on [bmj.com](#) gives full details of categorisation (according to that shown in the box) and characteristics. Eight studies were prospective, although the data extraction was primarily cross sectional. The remaining articles were cross sectional or retrospective surveys. Two studies were interventional: one a case-control study involving onsite training and the other a study carried out before and after a software update. Both showed substantial improvements in recording levels after the intervention. A retrospective cohort study also showed an increase in completeness and accuracy of EPRs over five years.

Structured data (codes, classifications, and nomenclatures) were most commonly investigated. Although textual data were mentioned, only one study considered it in any detail. Twelve documents did not present their data structure (that is, coding system name) while most did not present the precise codes being investigated. UK publications generally used Read and OXMIS (Oxford medical information systems) codes. In other countries the ICPC (international classification of primary care) codes were more widely used. ICD (international classification of diseases) codes act as a referencing standard for these primary care coding systems. Where necessary, subsidiary codes (for example, chapter headings from *British National Formulary*; Prescription Pricing Authority) were used.

Quality of data (reliability) was usually measured with rate comparisons. Data validity was expressed under a range of terms (completeness, correctness,

accuracy, consistency, and appropriateness), which were rarely defined. Sensitivity (completeness) was the commonest such index (table, [webextra table B on bmj.com](#)).

Seven studies carried out questionnaire and telephone surveying for a reference standard, one study used video recording, and 24 used clinical information gathered during the consultation (table). Seventeen publications used triangulation within the EPR to test internal consistency of data. Medication data was the most common internal reference standard. EPR diagnostic status was appraised through electronic prescribing information and subsequently validated against the paper notes. Hospital discharge details have also been used to evaluate EPR diagnostic status. The presence of hospital diagnosis and procedural data have been found to improve the quality of data in primary care. Eighteen studies used national statistics or survey data as a reference standard for data reliability. A third of UK studies used the fourth national study of morbidity in general practice (MSGP4).

Scoping data

Fifteen studies used EPR data for research or practice management. Although the intention of these studies was not to measure data quality, they gave insight into issues of data validation. These studies relied more on positive predictive value as a quality measure than sensitivity (table). Fourteen studies considered the diagnostic status of the patient, with 10 publications dealing primarily with information on patient identification and case validation. Three used survey techniques to establish diagnostic status. Of the 12 retrospective investigations, seven used centralised datasets. These "scoping" studies were more than twice as likely to present confidence intervals than studies that measured data quality (10/15 (67%) *v* 11/37 (30%)).

Levels of data recording

Prescribing data are generally the most sensitive, and the ability to link prescriptions with diagnosis was the favoured means of identifying patients and establishing the predictive validity of diagnostic codes. The sensitivity of other EPR elements was wide ranging, while positive predictive value was consistently high. Diseases with clear diagnostic criteria were generally better recorded, as were data on specific procedures. Lifestyle and socioeconomic data were rarely studied and then only in terms of sensitivity.

Discussion

We believe this is the first systematic review to investigate the measurement of quality of data in primary care. Most research has been published since 1995, reflecting the increasing importance and use of EPRs. Publications were mostly descriptive. This is indicative of a topic in which the direction of change is externally imposed (that is, controlled by the pace of technological development). The appraisal of data quality has favoured practices that embrace technology and so will be an overestimate of the general picture.

The dominance of UK publications is unsurprising given the scope of this review. This also suggests an understanding of the importance of the quality of EPR

Proportions of data type being investigated, reference standards used to assess quality, and commonest measures of quality. Figures are numbers (percentage) of studies

	Measuring data quality (n=37)	Scoping data quality (n=15)
Data type:		
Diagnostic or symptom state data	34 (92)	14 (93)
Management data	16 (43)	4 (27)
Wider aspect of care	10 (43)	3 (20)
Reference standard		
Video, direct examination:	1 (2)	1 (7)
Surveys	7 (19)	3 (20)
Routine data	24 (65)	12 (80)
National statistics, rates	18 (49)	8 (53)
Quality measures used:		
Rate comparison	27 (73)	10 (67)
Sensitivity	20 (54)	7 (47)
Positive predictive value	7 (19)	8 (53)
Specificity	3 (8)	1 (7)

data in terms of health policy and validated research databases, within the United Kingdom. Publications from non-English speaking countries were disadvantaged under our selection criteria. Those that were identified used similar techniques to measure data quality.

Measuring quality

The element of the EPR being investigated (numerator) and the components of the reference standard used to appraise its quality (denominator) were often not clearly defined within the literature (for instance, diagnostic code/diagnostic criteria). When they were defined there was inconsistency between studies. This makes comparisons risky and meta-analytical interpretation of results impossible.

Measurement theory requires that both the concepts of validity and reliability be addressed. Reliability (a precursor to validity) is a measure of stability and is appraised through the subjective comparison of rates and prevalence. Sensitivity and positive predictive value, the most widespread measures of data validity, presuppose that the selected denominator is an adequate representation of the true dimension being measured. To identify the real health status of the patient subjective (perceived), objective, and diagnostic dimensions need to be measured by different techniques and their appropriateness for EPR validation considered. To aid interpretation and make comparisons between populations, confidence intervals should be provided.

When the opportunity to record clinical data in different forms (paper and computer) exists, this decreases validity of any one to act as a true reference standard. The use of paper notes to assess EPR validity will become increasingly inaccurate as clinicians migrate to electronic systems. In the medium term it is best to consider several independent markers of quality, and those studies that used several explicit reference standards (triangulation) were more likely to reflect the true quality of electronic data (see table B on bmj.com).

To facilitate comparisons of data quality across sites and systems, it is essential to have a reference standard. In the longer term we recommend the establishment of internal reference standards based on those objective and diagnostic EPR elements recognised as

What is already known on this topic

The demonstration of quality is central to the NHS strategic agenda

Data from electronic records are expected to have a central role within healthcare commissioning, quality control, clinical governance, and the new GP contract

No standard methods of measuring data quality have been described

What this study adds

A framework for categorising and selecting papers which report data quality in primary care

Reliability of data was measured through rate comparison in 73% of studies, while validity was calculated mostly through measures of sensitivity

Markers of quality should comprise internal reference standards based on objective and diagnostic EPR elements that have high positive predictive value

having high positive predictive value (that is, diagnostic codes, prescriptions, test results, referral outcomes, procedural codes). Such reference standards can then be used to explore measures of sensitivity.

We thank E Mitchell (Tayside Centre for General Practice) and N Booth (Sowerby Centre for Health Informatics at Newcastle) for access to their specialist medical informatics databases.

Contributors: See bmj.com

Funding: Fisher Medical Centre Research Unit is funded by the NHS Executive Northern and Yorkshire Region. The guarantor accepts full responsibility for the conduct of the study, had access to the data, and controlled the decision to publish.

Competing interests: None declared.

- 1 Moss F. Spreading the word: information for quality. *Quality Health Care* 1994;46-50.
- 2 House of Lords. *Select committee on science and technology: fourth report*. 2001. <http://www.parliament.the-stationery-office.co.uk/pa/ld200001/ldselect/ldstech/57/5701.htm> (accessed 1 Jan 2003).
- 3 Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. *J Am Med Inform Assoc* 1997;4:342-55. (Accepted 6 February 2003)

One hundred years ago

The massacre of the innocents

Fresh evidence of the necessity of drastic legislation for the protection of children placed out to nurse is afforded by the tale told at the Central Criminal Court last week. Three women were convicted of serious crimes against children committed to their so called "care."

In the first case—that of Gale—the prisoner was sentenced to two years' hard labour; the woman pleaded guilty to having wilfully abandoned in railway trains and stations three children that certain unmarried women had confided to her for adoption, paying with them sums varying from £15 to £20. In the second case two women—Walters and Sach, a nurse—were convicted before the same court of the murder of a male child. The prisoner Sach kept a "maternity home," one of the advantages of which, as claimed in the advertisements, was that "baby can

remain." As the woman charged her cases 3 guineas a week, it is evident that her *clientèle* did not consist of the very poor class; her charge for having the baby adopted by "a lady" was £25. Walters appears to have been the woman to get rid of the children, as she was found with one wrapped up as a bundle in her arms dead; and it was proved in evidence that the same thing had occurred before with the same woman. Both were sentenced to death. No fewer than 300 articles of infants' clothing were found in the prisoner's house. One of the witnesses stated that during the time she was there no fewer than twenty women were confined in the house. The fact that about a hundred children's dead bodies are found in London alone by the police every year proves that infanticide is of very common occurrence if conviction for the crime is not. (BMJ 1903;i:221)