Family Practice, 2014, Vol. 31, No. 2, 127–129
doi:10.1093/fampra/cmu009
Advance Access publication 13 March 2014

Editorial

Electronic medical records: the way forward for primary care research?

Introduction

Electronic medical records (EMRs) are becoming the norm in many health systems internationally, especially in the primary care setting. Though designed to help family doctors and other clinicians to record and manage patient care more accurately and efficiently, they are often useful for research purposes too. Indeed recent years have seen huge advances in the quality, availability and use of EMR databases for research.

This increased use of EMRs for research has led to work, such as a recent paper from The Netherlands, attempting to establish quality criteria for these EMRs to be used in research (1). In the UK, the General Practice Research Database has recently become the Clinical Practice Research Datalink (CPRD) and aims to substantially extend its coverage in terms of population size and also the sources of data available (2). As with several Scandinavian registries [e.g. (3)], CPRD data can be linked with national registers (e.g. mortality, cancer), as well as sociodemographic and hospital admissions data. Until recently, the majority of primary care EMRs suitable for use in research have been from Western European countries, possibly due to their health care systems readily facilitating this sort of data collation. This is now changing, with a notable example of an up-and-coming EMR for use in research being Canadian Primary Care Sentinel Surveillance Network (4).

Potential uses of EMRs in research

In an editorial in this journal in 2012, Martin Dawes (5) described a mismatch between the conditions making up the primary care workload, and how well this is reflected in the topics of published primary care research. The mismatch he described might be as a result of researchers not fully understanding at a quantitative level what real-world primary care looks like. EMRs can provide an overview of the true make-up of primary care practice workload (3), as well as sufficient numbers for a study that might be difficult (e.g. relatively rare disease) if primary data collection were required (6,7). EMRs also afford the possibility to study events that are otherwise difficult to capture. For example, in a recent issue of this journal, Willems *et al.* (8) used a Dutch database to study benzodiazepine doses. They refuted the widely held belief that doses needed to be increased over time in long-term users. Without the use of routinely collected data, this study would have been near on impossible, due to the social acceptability bias that would likely surround such a study. The description of actual consultation and prescribing habits is generally free of such biases, to which self-reported information can be prone. Additionally, the comprehensive nature of EMRs, coupled with large sample sizes and the ability to follow patients over long periods of time, allows for a wider range of variables to be considered (provided they have been recorded for clinical purposes). For example, in their study, recently published in this journal, Ursum *et al.* (9) were able to evaluate 121 co-morbidities in those newly diagnosed with inflammatory arthritis. This would have been very difficult in a primary research study, without linking study data to clinical records. This linkage is of course another use to which EMRs have been put in research studies [e.g. (10)].

Potential pitfalls of EMRs for researchers

Despite the advantages of using EMRs for research purposes, there are a number of drawbacks, and these often appear to be ignored by authors. First and foremost, EMRs are created for clinical and not research purposes. This means that although some aspects of health care are likely to be very comprehensive, for example in the UK all primary care prescriptions should be recorded electronically, the same is not necessarily true for other aspects of care. The record of symptoms and diagnoses is a combination of what was presented to the doctor by the patient, and then what the doctor chose to record. It may not give a full picture of the patient's situation. Furthermore, some variables that would be routinely collected in a research study may never be entered into an EMR. For example, studies of pain would usually include a measure of pain severity, but this is unlikely to be entered into an EMR, and if it is, it will likely prove difficult to extract this information in a systematic way. Similarly, much

information may be hidden in the 'free text' of consultations, and while work is ongoing to harness this data [e.g. (11)], it is far from being available on a routine basis at the present.

A major criticism from peer reviewers of papers using EMR data is the potential for inaccuracies in diagnosis. This raises the crucial issue of understanding the context of EMR data, which varies from database to database and from study to study. Researchers using EMRs should be aware of when a particular symptom or diagnosis is usually entered into the record in that database. This is likely the diagnosis that the clinician made at the time and in some studies will be of direct interest. However, in other studies, researchers may need to consider how they might 'validate' a diagnosis. Examples of this might include using a published algorithm to define the diagnosis of interest (9), or ensuring those with a coded diagnosis also have a relevant prescription. Again, when using prescription data, it is important to consider how the health system from which the EMR is extracted might influence prescribing behaviour, as well as the use of prescribed medicines by the patient. In England, for example some patients are required to pay a flat fee for any prescribed drug, but others (e.g. children, the elderly, those on low incomes) are not. So for drugs available without a prescription, the doctor may recommend a particular treatment (e.g. paracetamol/acetaminophen) to some patients without writing a prescription, while others receive the prescription.

A final and often underappreciated drawback to using EMRs for research is the computing power and skill required to make use of these clinical records for another purpose. This necessitates access to appropriate hardware, as well as software and appropriately skilled staff, and should not be underestimated.

Future uses of EMRs

A wide range of pharmacoepidemiological and other observational studies have been undertaken in EMRs, and there is now a move towards embedding randomized clinical trials within databases. In these studies, patients are randomized at the point of care and high rates of follow-up are more-or-less guaranteed, as outcomes are captured entirely within the EMR (12). Furthermore, as the range of available data in EMRs increases, more types of studies will be possible. For example, Wynne-Jones *et al.* (13) used a local database to assess rates of sickness certification in the UK, a study that would have been problematic without EMRs.

Summary

While EMRs present a potentially powerful research tool, those considering their use for research should not be complacent about the amount of work and skill involved. A large amount of computing ability is often required, alongside the need to fully understand the context of the data, as well as the overarching clinical question.

Acknowledgements

The author would like to thank Dr Richard Hayward and Dr Kelvin Jordan for their comments on a previous version of the manuscript.

Declaration

Funding: National Institute for Health Research School for Primary Care Research to SM. This article presents independent research funded by the National Institute for Health Research. The views expressed are those of the author(s) and not necessarily those of the National Health Service, the National Institute for Health Research or the Department of Health.
Ethical approval: none.
Conflict of interest: none.

*Sara Muller**
Research Institute for Primary Care & Health Sciences, Keele University, Keele, UK.
**Correspondence to Sara Muller, Research Institute for Primary Care & Health Sciences, Keele University, Keele ST5 5BG, UK. E-mail: s.muller@keele.ac.uk*

References

1. van den Dungen C, Hoeymans N, Schellevis FG, van Oers HJ. Quality aspects of Dutch general practice-based data: a conceptual approach. *Fam Pract* 2013; **30**: 355–61.
2. Clinical Practice Research Datalink. CPRD.com (accessed on 10 March 2014).
3. Wändell P, Carlsson AC, Wettermark B, Lord G, Cars T, Ljunggren G. Most common diseases diagnosed in primary care in Stockholm, Sweden, in 2011. *Fam Pract* 2013; **30**: 506–13.
4. Canadian Primary Care Sentinel Surveillance Network. cpcssn.ca (accessed on 10 March 2014).
5. Dawes M. Symptoms, reasons for encounter and diagnoses. Family practice is an international discipline. *Fam Pract* 2012; **29**: 243–4.
6. Hamoen EH, Reukers DF, Numans ME, Barentsz JO, Witjes JA, Rovers MM. Discrepancies between guidelines and clinical practice regarding prostate-specific antigen testing. *Fam Pract* 2013; **30**: 648–54.
7. Muller S, Hider SL, Belcher J, Helliwell T, Mallen CD. Is cancer associated with polymyalgia rheumatica? A cohort study in the General Practice Research Database. *Ann Rheum Dis* 2013 July 10 [Epub ahead of print] doi:10.1136/annrheumdis-2013-203465.
8. Willems IA, Gorgels WJ, Oude Voshaar RC, Mulder J, Lucassen PL. Tolerance to benzodiazepines among long-term users in primary care. *Fam Pract* 2013; **30**: 404–10.
9. Ursum J, Korevaar JC, Twisk JW *et al.* Prevalence of chronic diseases at the onset of inflammatory arthritis: a population-based study. *Fam Pract* 2013; **30**: 615–20.
10. Barber J, Muller S, Whitehurst T, Hay E. Measuring morbidity: self-report or health care records? *Fam Pract* 2010; **27**: 25–30.
11. Wang Z, Shah AD, Tate AR, Denaxas S, Shawe-Taylor J, Hemingway H. Extracting diagnoses and investigation results from unstructured

text in electronic health records by semi-supervised machine learning. *PLoS One* 2012; **7**: e30412.

12. Gulliford MC, van Staa T, McDermott L *et al.*; electronic Cluster Randomised Trial Research Team eCRT Research Team. Cluster randomised trial in the General Practice Research Database: 1. Electronic decision support to reduce antibiotic prescribing in primary care (eCRT study). *Trials* 2011; **12**: 115.
13. Wynne-Jones G, Mallen CD, Mottram S, Main CJ, Dunn KM. Identification of UK sickness certification rates, standardised for age and sex. *Br J Gen Pract* 2009; **59**: 510–6.