Electronic Health Records Research in a Health Sector Environment with Multiple Provider Types

Samantha S.R. Crossfield, Susan E. Clamp
Leeds Institute for Health Sciences, University of Leeds, 101 Clarendon Road, Leeds, UK
{s.crossfield, s.clamp}@leeds.ac.uk

Keywords: Research Database, Shared Health Records Access, EPR, Patient Record Access, ResearchOne, Multiple Health Providers, Clinical System, Database Validation.

Abstract: Where healthcare provision is divided into provider types, such as child health and palliative care, it is difficult for researchers to access comprehensive healthcare data. Integrated electronic health records offer an opportunity for cross-provider type care research. In this paper a new model for accessing such data is justified using the critical success factors as determined from an established research data provider. This validates a model that will enhance integrated health research for the benefit of clinical practice across multiple provider types.

1 INTRODUCTION

In gaining health and social data for research, and transferring results into clinical practice, the consideration of multiple health provider types that are treating a patient is often relevant and yet receives little attention. Electronic health record (EHR) data is increasingly being used in research due to their widespread use in clinical practice for gathering detailed and structured data. EHRs are often non-shareable, used in only the health unit that generated them, such as a general practice or a hospital ward (ISO/TR 20514, 2004). Such records cannot comprehensively represent to the research community the health of patients who receive care in multiple settings. Alternative structures with EHR data sharing between clinicians at different health units can improve clinical practice and reduce errors (Twomey et al., 2004; Ammenworth et al., 2008). Research on integrated EHRs has enhanced healthcare through investigating clinical practice through a whole systems approach. Research using integrated EHRs in England has nonetheless remained infrequent due to accessibility issues.

Health care in many countries including Spain, UK and Germany is delivered through multiple provider types working with independence. These care providers struggle with an outcome of this specialisation which is often termed ‘silo working’ wherein service deliverers with different aims and professional languages gather information on separate aspects of patient care (Wilson et al., 2007; Kawonga et al., 2012) and store these in unrelated, closed silos. These records may not be shared with other health units, let alone across provider types. As a consequence each unit holds partial patient records rather than the entirety of the patient’s medical history. With closed systems it becomes difficult to share timely and pertinent information, such as diagnoses, allergies, medication and professional insights with other healthcare providers that are also intervening with and monitoring the health of the patient. This results in issues of duplication and missing data. Patient information held in such silos provides less support to patients that cross care provider types and reduces the capacity to perform longitudinal assessments (Kuperman, 2011).

This silo issue is also of relevance for the research community who consider patient health in an array of fields including health informatics, epidemiology, health economics, clinical care and medicine. Traditional data collection involves such invasive, timely and resource-intensive methods as conducting interviews and questionnaires. The increasingly routine use of EHRs in clinical practice, including by 76% of European general practitioners (Dobrey et al., 2008), makes EHRs an efficient source of large cohort research data. The capacity
for large EHR cohorts facilitates research on low frequency incidences or diagnoses. This enabled identification of the correlation between emergency department waiting times and outcomes of mortality and readmission, identified using the records of over 14.5 million emergency department attendances (Guttmann et al., 2011). However patient data dispersal in EHRs across multiple provider types necessitates the requirement of identifiable data for undertaking data linkage. This brings security issues and the time taken to gather and link siloed data reduces the timeliness of cross-provider research. Further time on the part of the researchers and the data providers is required to update the research dataset. The ethical issues surrounding the identification of relevant patients and in developing a fully informed consent mechanism remain. Nevertheless such research has successful results and was crucial in resolving the Autism-MMR vaccine dispute (Taylor et al., 1999). In using non-shared EHR data researchers face the same constraints as clinicians, in not being able to view the full patient pathway in a timely, cost-effective and secure, audited manner.

In May 2012 the Department of Health in England issued a call for efficient EHR research. QResearch has an established ten year record in supporting this. QResearch is a UK-based not-for-profit general practice EHR research database. It was developed with the aim of consolidating de-identified, siloed EHR data from a large representative cohort of general practices, with the aim of providing data for ethical research purposes (Hippisley-Cox et al., 2004). It has facilitated such research as the development of tools for identifying patients at risk of developing cardiovascular disease (CVD) and diabetes (Hippisley-Cox et al., 2008; Hippisley-Cox et al., 2009). Such developments using non-comprehensive, non-shared records may have limited external validity and global relevance both within and beyond general practice. This may influence their usage in clinical practice as 66% of clinicians who identify the need to perform a global CVD risk assessment fail to follow the guidelines by employing such a tool, valuing subjective assessment alone (Graham et al., 2006). Cross-provider EHR research will be required in answering this call issued by the Department of Health in a health sector environment with multiple care provider types.

There are well established alternatives to non-integrated clinical practice and research. Kaiser Permanente provides comprehensive care packages to 8 million patients in America, with shared standards between care providers. In the UK SystmOne is a centrally hosted clinical system provided by TPP that enables record sharing between many of the health provider types in the National Health Service (NHS) (Figure 1). For 15 years its centralised database has contained one integrated EHR per patient. From this, data is shared with the patient and across the health units that use SystmOne, where access rights are legitimate. Through SystmOne 25 million patients have a shareable record (Table 1) that facilitates integrated care (Stoves et al., 2010). Both Kaiser Permanente and SystmOne exemplify long-standing alternatives that reduce the ‘silo effect’ in healthcare.

---

Figure 1: Healthcare provider types in which SystmOne is used and between which information can be shared with the exception of prison data. In hospitals SystmOne provides patient administration, clinical record viewing, bed management, accident and emergency, e-prescribing and e-discharge.
Table 1: Approximate count of units and patients with a relationship recorded on SystmOne from ten care provider types. Other provider types using SystmOne include Speech and Occupational Therapies, Community and Social Services, Dietetics, Palliative Care, School Nurses and Endocrinology.

<table>
<thead>
<tr>
<th>Health care provider type</th>
<th>Patients with data on SystmOne</th>
<th>Count of units using SystmOne</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Practice</td>
<td>19 million</td>
<td>1900</td>
</tr>
<tr>
<td>Child Health</td>
<td>6 million</td>
<td>50</td>
</tr>
<tr>
<td>District Nursing</td>
<td>5 million</td>
<td>1250</td>
</tr>
<tr>
<td>Health Visitor</td>
<td>2.4 million</td>
<td>180</td>
</tr>
<tr>
<td>Out of Hours</td>
<td>2.9 million</td>
<td>120</td>
</tr>
<tr>
<td>Physiotherapy</td>
<td>1.3 million</td>
<td>60</td>
</tr>
<tr>
<td>Acute Hospital</td>
<td>1.1 million</td>
<td>10</td>
</tr>
<tr>
<td>Podiatry</td>
<td>1.1 million</td>
<td>40</td>
</tr>
<tr>
<td>Community Primary Care Clinic</td>
<td>948,000</td>
<td>270</td>
</tr>
<tr>
<td>Minor Injuries / Accident and Emergency</td>
<td>864,000</td>
<td>50</td>
</tr>
</tbody>
</table>

Integrated EHRs assist in cross-provider type care management that efficiently utilises resources. Integration at Kaiser Permanente units contributes to the number of bed stays being 3.5 times fewer than in the NHS for 11 leading causes (Ham et al., 2003). Benefits from EHR sharing are indicated by patient management improvements in cases that involve professionals from primary and secondary care sectors, such as are frequent in the treatment of long-term conditions. In the case of diabetes management this is evidenced by, for example, secondary care consultants referring more frequently to an EHR where the patient’s general practitioner (GP) is also registered with SystmOne as the record would contain updates since the previous outpatient appointment and contains the advice of other specialists, recent medications and blood test results (Keen and Denby, 2009). This replaces reliance upon physical meetings, patient awareness or being able to telephone other care providers (Stoves et al., 2010). Comprehensive information allows consultants to review and communicate medication modifications, while enabling primary care nurses to monitor ongoing treatments and identify where patients have non-attendance at other services (Keen and Denby, 2009). Shared records also assist in prompt medicine reconciliation between care settings, identifying errors in 38% of prescriptions (Moore et al., 2011). Clinical system integration delivers the benefits of an “electronic highway” envisioned by the NHS National Programme for IT (Department of Health, 2003).

The benefits that integrated records bring to clinical practice, in terms of the timely provision of comprehensive information-sharing, could also be brought to the research community. Integrated EHR Kaiser data supports research that considers care provision across multiple provider types. Clinical practice has altered internationally in response to links uncovered between hospital admissions and drugs such as Vioxx being issued in ambulatory (primary) care settings (Graham et al., 2005; Cheetham et al., 2009). Using shared EHRs in research replaces linkage exercises that involve identifiable data and result in biased, incomplete datasets (Bohensky et al., 2010). Shared EHRs enable research on the otherwise lost communications between care providers, such as referral trails. Research on cross-provider type records could validate siloed research in a cost-effective, timely manner and inform clinical practice that occurs in integrated settings.

The aim of this paper is to determine the capacity of a new research database, ResearchOne, to facilitate cross-provider EHR research in the UK. ResearchOne is a not-for-profit organisation with ethical approval to extract de-identified EHR data from the centralised, cross-provider SystmOne database into ResearchOne. From here secure, audited access to records data by the research community has the purpose of developing research that improves healthcare. This model must be investigated in order to justify that ResearchOne may bring benefits to research in the way that SystmOne does for clinical practice.

2 METHOD

The method was designed to assess the capacity of the ResearchOne database to support EHR research and to justify its potential benefits to integrated records research using English health data. Information regarding SystmOne and ResearchOne were determined from the ResearchOne Database Protocol and through interviews (Crossfield et al., 2012). QResearch is specifically designed for EHR research in the UK and was taken as an academically established ‘standard’. The key features of QResearch were taken as critical success factors, as justified in Table 2, against which the model presented by the ResearchOne database was appraised. These factors are the headings in the following section. From this ResearchOne could be validated with the potential to perform to the
existing standard for a research database of NHS data, in order that it can facilitate cross-provider research.

Table 2: Features of the QResearch general practice EHR research database, with reasoning behind their necessity.

<table>
<thead>
<tr>
<th>Critical factor</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data consolidation</td>
<td>The database assists researchers in accessing data that has been consolidated from many health units and so reduces the invasion, time and cost for clinicians and researchers, who must otherwise perform repeated extracts.</td>
</tr>
<tr>
<td>Large cohort of research EHR data</td>
<td>Larger sample sizes bring both power and validity to research outcomes, enabling more research questions to be addressed (Cohen, 1992).</td>
</tr>
<tr>
<td>De-identified EHR data</td>
<td>De-identification of EHR data protects privacy and permits research access without a public health mandate or consent, which could not be feasibly and non-invasively acquired for a significantly large cohort (Lowrance, 2003; Wellcome Trust, 2009).</td>
</tr>
<tr>
<td>Representative coverage</td>
<td>The external validity of a research outcome depends upon it being derived from a representative sample of the population.</td>
</tr>
<tr>
<td>Ethical research practice</td>
<td>Success relies upon the database being securely developed and used for ethical purposes.</td>
</tr>
</tbody>
</table>

3 RESULTS

3.1 Data Consolidation

QResearch facilitates research on EHR data consolidated from over 650 non-integrated general practice (GP) databases (Vinogradova et al., 2012). The ResearchOne database can similarly hold EHR data contributed by multiple practices. Moreover it can hold data from other unit types, as it mirrors the successful infrastructure of SystmOne and so it meets this critical success factor. SystmOne integrates data from multiple units into one centralised record per patient and so no consolidation is required in order to extract data in SystmOne from multiple provider settings into the ResearchOne database.

Data linkages to other sources undertaken by QResearch are also feasible with ResearchOne. QResearch links GP EHR information to socioeconomic, Hospital Episode Statistics (HES), disease-specific registry and death registration data (Hippisley-Cox et al., 2008; Hippisley-Cox and Stables, 2011). ResearchOne has national ethical and governance approval to perform such linkages and will likewise consolidate such information (Crossfield et al., 2012). An NHS National Institute for Health Research funded study, Improving Prevention of Vascular Events in Primary Care, has successfully piloted the capacity to link ResearchOne data to HES and Myocardial Ischaemia National Audit Project (MINAP) data.

3.2 Large Cohort of Research EHR Data

General practice involvement in QResearch has grown steadily to over 650, surpassing the original aim of 500 practices (Hippisley-Cox et al, 2004; Vinogradova et al., 2012). SystmOne hosts patient information for more than 25 million patients across England from more than 4500 units that may participate in ResearchOne. This includes 1900 general practices, 110 community services and 50 palliative care units (Table 1). Prison data recorded in SystmOne cannot be extracted into the ResearchOne database. TPP already hosts this data in SystmOne and has the data management skills and capacity to hold such a large cohort of records from multiple provider types in ResearchOne.

3.3 De-Identified EHR Data

Both the ResearchOne database and QResearch have a nationally approved governance framework under which they can hold only de-identified data. Neither database can contain free text with potentially identifiable data, nor full dates of birth or death. Furthermore, given the comprehensiveness of cross-provider type records, the ResearchOne database excludes diagnostic cases that are present in fewer than five records. QResearch requires consent from each practice in order to access their databases to perform the data extraction. While SystmOne is centrally hosted, ResearchOne follows this practice in requesting consent from contributing health units, and also provides the opportunity for patients to ‘opt out’. Consent is electronically audited through SystmOne, the centralisation of which ensures that any changes will automatically update ResearchOne within seven days.
3.4 Representative Coverage

QResearch practices are “spread throughout the UK”, offering representative general practice coverage (Hippisley-Cox et al., 2004, p.49). ResearchOne has the capacity to provide an England-wide representation of cross-provider type healthcare, through the more than 4500 invited units. England is divided into 433 lower tiers of local government – Local Authorities - of which over 85% have patient representation on SystmOne. There are more than 25 million patients contributing to 300 million years of patient records, geographically distributed across England. Of these patients 13 million are registered with more than one care unit on SystmOne, and 118,000 patients are actively receiving care from five or more units that use SystmOne. SystmOne holds 4.8 billion diagnostic codes, inputted by clinicians whose specialties range from ante-natal to geriatric, rehabilitation to neuropathology. This coverage is across community, primary and secondary care. The capacity for representation also covers the indices for rurality and deprivation defined by the UK Economic and Social Data Service (2012).

3.5 Ethical Research Practice

QResearch and ResearchOne are specifically designed for ethical research access with the aim of improving healthcare. The frameworks for both QResearch and ResearchOne have been developed with ethical and governance approval from the relevant national bodies (Hippisley-Cox and Stables, 2011; Crossfield et al., 2012). Any change in policy is reviewed by the national boards and a database committee of patients and clinical professionals along with experts in informatics, database architecture and governance. ResearchOne and QResearch have an approved framework to review data requests based on the benefit of project proposals to clinical practice and whether they will produce publishable results (Hippisley-Cox and Stables, 2011; Crossfield et al., 2012). Ethical accessibility of both QResearch and ResearchOne is supported by their being not-for-profit organisations. With SystmOne data existing centrally, the cost of ResearchOne maintenance is low, which reduces the cost further for the research community. Remote access to the secure ResearchOne data warehouse is audited for the purpose of maintaining this ethical practice.

4 DISCUSSION

The results of this investigation show that ResearchOne matches the capacity of the existing standard of a research database with EHR data. ResearchOne has been nationally approved to extract de-identified EHR data from consenting health units. There is potential for the inclusion of data from a large cohort of shared EHRs with representative coverage both geographically and demographically across England. The framework has been designed to ethically support research that delivers benefit to patient health.

4.1 Further Research Capacities

ResearchOne has the capacity to perform beyond this standard. The ResearchOne database has the potential to maintain data from multiple units across more provider types in England than the current standard, and so can support more comprehensive and representative research between and within these areas of healthcare. As this data is integrated via SystmOne, no biased and potentially unsecure linkage exercises are required. This is beneficial because Bohensky et al. (2010) reviewed linkage sensitivity to range between 74-98%. The capacity to extract data from centralised records moves the ResearchOne database beyond the current standard of consolidating data from isolated units. This does not disturb clinical practice and reduces the cost of extraction, with this saving being passed to the research community. The centralisation of SystmOne maintains an up-to-date audit of unit consent and also enables a patient to opt out of the ResearchOne database by informing just one of their care providers. Such further capacity is of relevance in ethically supporting the research community to enhance healthcare.

A further factor that assists in health research is the timeliness of the data used, and EHRs can and should provide timely information (Powell and Buchan, 2005). Timely data is required for research in order to reflect the evolving field of clinical practice in a country with changing population and health demographics. The centralisation of SystmOne ensures that research data could reflect real-time clinical developments, without affecting SystmOne users. The ResearchOne database framework involves updates from SystmOne at least every seven days; this could not be compared to the QResearch full database update frequencies that could not be found to be listed. Items can be
extracted more frequently should this be required for public health surveillance purposes. A further reason for timeliness is that consent withdrawal from the ResearchOne database results in their data being removed within seven days. Timely data provision can occur securely at minimum cost due to the centralised nature of SystmOne, which is a more speedy and accurate alternative than linking data from multiple provider sources.

ResearchOne has the capacity to bring the benefits of record-sharing into the research arena. SystmOne data that is extracted into ResearchOne is not consolidated from multiple settings, but rather it is integrated between them. Shared EHRs can record data that isolated records cannot, even when linked. Through ResearchOne the integrated care records of more than 25 million patients have the capacity to support research, with information from over 4500 health units in primary, secondary and social care (Table 1). With integrated EHRs there is less missing data that could impact on research validity and similarly no duplication of data if, for example, a patient moves to a new unit on the same clinical system. This presents a more comprehensive view of the health sector and the delivery of patient care.

4.2 Next Steps

The capacity of ResearchOne to maintain a large cohort of de-identified EHR data from multiple provider types depends upon unit participation. The joining process is simple and extracts will not inconvenience SystmOne users given its centralised nature. Data providers can be assured that the ResearchOne database is maintained under the same security principles as SystmOne in an NHS-accredited data centre. The aim of ResearchOne, to relay research outcomes back into clinical practice, assures that it is beneficial to contribute to the ethically approved process. The success of QRResearch should assist SystmOne users, some of whom may have contributed to QRResearch previously, in recognising this beneficial invitation. The opportunity to participate in ResearchOne has begun to be rolled out to health units on SystmOne and is successfully indicating the realisation of the capacity of ResearchOne.

The aim of ResearchOne includes not only pulling data for research purposes, but pushing outcomes back into healthcare. Results must be openly published, and where these have relevance to clinical care across many provider types, they can initiate more comprehensive healthcare benefits. SystmOne will incorporate developments so that the clinical system continually improves the support it provides to over 120,000 clinicians across multiple provider types. This cyclical, evaluative model of clinical practice and research that is encapsulated in SystmOne and ResearchOne can be envisioned as a global model for the future.

ResearchOne will facilitate research both for validation purposes and in novel areas. Explorations will compare ResearchOne data with national statistics to validate its representative coverage. Research performed on other datasets will be validated using the ResearchOne database. The impact of integrated and isolated EHR data on research will also be investigated, to explore the role of ResearchOne. Research projects may use data from single or multiple care provider types, as well as national statistics and registry data that are incorporated into the database. In this way ResearchOne aims to facilitate results that are of relevance across all units that are contributing data.

The global capacity for EHR research is continually increasing. Progression in science, particularly in the fields of security technology and machine learning, will lead to data mining of entire EHRs. The anonymisation of free text through advancements in natural language programming, and the reduction of human involvement in data analysis will open up the capacities of EHRs to ethically support research. The number and types of units across which SystmOne provides integrated EHRs is increasing, with a recent number of acute deployments being made, so that ResearchOne has increasing potential to represent comprehensive care in England. With future developments ResearchOne will increasingly support research that benefits healthcare.

5 CONCLUSION

The contribution of EHRs to research is increasing, but is hindered by the division of data across healthcare units as a result of the ‘silo effect’ in clinical care. Benefits in integrated care delivery have been evidenced from record-sharing. ResearchOne offers an alternative for research using shared EHR data. The model of the ResearchOne database has been critiqued using the success factors of QRResearch, an established provider of EHR
research data in the UK. ResearchOne meets this existing standard and brings further developments to the research community, especially in terms of the timely provision of integrated, cross-provider type data and in feeding results back into clinical care. This offers a global model for integrated evolution between clinicians, patients and research.

REFERENCES


Economic and Social Data Service. [Accessed on 28.08.201] URL: hww.ccsr.ac.uk/esds/variables/


