Views of healthcare professionals to linkage of routinely collected healthcare data: a systematic literature review

Y M Hopf, C Bond, J Francis, J Haughney, P J Helms

ABSTRACT

Objective To review the literature on the views of healthcare professionals to the linkage of healthcare data and to identify any potential barriers and/or facilitators to participation in a data linkage system.

Methods Published papers describing the views of healthcare professionals (HCPs) to data sharing and linkage were identified by searches of Medline, EMBASE, SCOPUS, CINAHL, and PsychINFO. The searches were limited to papers published in the English language from 2001 to 2011.

Results A total of 2917 titles were screened. From these, 18 papers describing the views of HCPs about data linkage or data sharing of routinely collected healthcare data at an individual patient level were included. Views were generally positive, and potential benefits were reported. Facilitators included having trust in the system including data governance, reliability, and feedback. Some negative views, identified as barriers were also expressed including costs, data governance, technical issues, and privacy concerns. Effects on the physician–patient relationship, and workload were also identified as deterrent.

Discussion From the published literature included in this review, the views of HCPs were in general positive towards data sharing for public health purposes. The identification of barriers to contributing to a data linkage system allows these to be addressed in a planned data linkage project for pharmacovigilance. The main barriers identified were concerns about costs, governance and interference with the prescriber–patient relationship. These would have to be addressed if healthcare professionals are to support a data linkage system to improve patient safety.

BACKGROUND AND SIGNIFICANCE

Pharmacovigilance is the process by which adverse effects of drugs are detected, assessed, understood, and prevented. However, current systems are generally considered to be suboptimal and new approaches are required. Existing systems of pharmacovigilance vary and may be unique to individual countries, but often use elements of signal generation or dedicated follow-up studies. The term ‘signal generation’ refers to strategies designed to identify potential causal relationships between drug exposure and an adverse event. The Yellow Card Scheme (YCS) in the UK relies on voluntary reporting of adverse drug reactions (ADRs) by healthcare professionals (HCPs) and, more recently, by patients. In adults, all suspected adverse reactions to newly licensed drugs (currently bearing an inverse black triangle in the British National Formulary) should be reported. For children, any suspected adverse drug reaction, independent of licensing status or severity of event, should be reported. Although the YCS is well established, its recognized limitations include the voluntary nature of reporting, the lack of a denominator (the total number of exposed individuals), the duration of therapy, and the variable quality of the data received.

Although pre-licensing clinical studies aim to identify possible ADRs in addition to their clinical effectiveness, the limited number of patients in these studies reduces the chance of identifying uncommon reactions. At licensing it is not uncommon for less than 2000 patients to have been exposed to the given drug, whereas as many as 30 000 exposed individuals may be required to identify an ADR with an incidence of 1:10 000. Establishing sufficiently large exposure cohorts can be challenging, particularly for orphan drugs or rare conditions. This is compounded, particularly in children, by the frequent use of off-label or unlicensed drugs which are not subject to rigorous post-marketing surveillance. Linkage at individual patient level of routinely acquired health data between primary and secondary care could be important; rates of off-label or unlicensed prescribing are higher in secondary (hospital) care, but current systems cannot link reactions reported to the general practitioner (GP) with the original hospital prescription.

Linkage of routine healthcare datasets by unique patient identifiers could provide an alternative or complementary approach to the identification of ADRs. It would permit following exposed individuals in real time and provide a denominator. Routine data linkage would also enable creation of exposure cohorts in order to monitor long-term outcomes and enable a more efficient screening for side-effects or ADRs due to an ever increasing data pool. The use of routine healthcare data in the identification of potential adverse reactions to medicines through signal generation as well as the investigations of associations between an adverse reaction and medicines has been described previously.

Whole population single databases compiled for administrative purposes such as the Hospital Episodes Statistics (HES) in England have demonstrated their potential in monitoring disease trends and important health outcomes, and in Sweden the use of linked national datasets for whole population epidemiology, or what is judged to be clinically relevant research, has been made possible without individual patient consent by the Swedish Health Act and European Union directive 95/46/EC.

Although data linkage is mostly seen as advantageous, in particular for pharmacovigilance, some concerns have been voiced regarding
confidentiality and data protection for patient identifiable data,\textsuperscript{18–20} as well as practical issues about incomplete or missing data in routinely collected datasets.\textsuperscript{12, 21, 22} A Wellcome report published in 2009 in the UK recognized the potential of the use of electronic records for data linkage and research but also highlighted that clinical data can rarely be anonymous, in the full sense of the word.\textsuperscript{23}

To the best of our knowledge there is no current system which uses routine linkage of healthcare data for the purpose of identifying and monitoring ADRs. The CHIMES (Child Medical Records for Safer Medicines) program in Scotland is a research project which is developing a new system for drug monitoring and surveillance based on a linkage of routinely collected healthcare data from primary and secondary care, and prescription data. The focus of this initiative will be on children as it is known that ADRs in this population are particularly highlighted that clinical data can rarely be anonymous, in the full sense of the word.\textsuperscript{23}

OBJECTIVE

The objective of this literature review was to address the following questions: (1) What are the perceived barriers and facilitators to the linkage of routinely acquired healthcare data from the perspective of HCPs? (2) Would data linkage of routinely acquired healthcare data for clinical and research purposes be acceptable to HCPs?

METHODS

A systematic approach was used to review the current literature (as shown in figure 1).

Bibliographic databases and search strategy

Medline (Ovid Medline In-process and other Non-Indexed Citations and Medline (R) 1984), EMBASE (Embase Classic and Embase 1947), SCOPUS, CINAHL, and PsychINFO were searched. Search terms were adjusted to match individual database criteria. Each search comprised three broad domains: (i) medical records and data linkage, (ii) different types of HCPs, and (iii) views and opinions. The full search strategy is available on request.

Study inclusion criteria

Peer-reviewed, empirical papers, and conference abstracts covering primary and secondary research were eligible for inclusion. Qualitative and quantitative studies were included. The search was restricted to papers published in English from 2001 to 2011.

HCPs eligible for inclusion were medical doctors, nurses, or pharmacists.

Papers were included if they reported on views of HCPs on data sharing (ie, the shared use of information about an individual patient across settings), or data linkage (ie, the secondary use of aggregated, merged data across settings) of healthcare data, including clinical, administrative, and prescribing information, for example, from primary to secondary care.

Study selection

Titles and abstracts were reviewed by YH for eligibility. Full articles were retrieved for assessment or further clarification, for example if no abstract was available.

Data collection process

A data abstraction form (available on request) was used to record standardized information from each paper as follows: authors, citation, design, aims and objectives of the study, methods, setting and participants (number if provided), the type of data linked, the purpose of data linkage/sharing, and a summary of the key findings on barriers and facilitators.

Quality assurance

A random sample of papers was discussed at fortnightly research team meetings to confirm inclusion or exclusion decisions. A further random selection of papers was reviewed in duplicate by the researcher (YH) with several weeks between assessments. Papers for which the initial reviewer (YH) was unsure about inclusion or exclusion (n=26) were discussed with a second member of the research team (CB). In 23 cases the second reviewer confirmed the initial decision of YH (exclusion for n=20 papers and inclusion for n=3). For the remaining three papers a decision was made after discussion (n=2 included, n=1 excluded).

RESULTS

Screening and identification of papers

The search identified 2917 unique titles for screening. Selection of abstracts and papers is detailed in figure 1. A total of 188 abstracts were reviewed. For seven papers no abstract was available and the full paper was reviewed. One hundred and fifty-six papers were eligible for full review. Two papers were unavailable, that is, could not be retrieved during the review time, leaving 154 full-text articles. Authors were contacted in six cases to ask for further information/clarification (two authors answered and supplied further information). Papers/abstracts were excluded due to the following reasons: not about data sharing/linking across settings (n=47), no views of HCPs to data sharing/linkage (n=35), not describing empirical research (n=28), and other reasons (n=26).

Study characteristics

An overview of the characteristics of the 18 included studies is presented in online supplementary table S1. The majority were conducted in the USA,\textsuperscript{26–34} followed by Canada,\textsuperscript{35–39} Two studies were conducted in each of the UK,\textsuperscript{38} 40 and the Netherlands,\textsuperscript{41, 42} and one study in Finland.\textsuperscript{43}

Many studies used surveys to explore views,\textsuperscript{28, 30, 32, 34, 36, 41, 43 six studies used qualitative methods such as interviews and focus groups,\textsuperscript{27, 28, 35, 39, 41, 42} and the remaining five used a mixture of both and were classified as mixed methods research.\textsuperscript{26, 31, 33, 37, 40}

Several sampling methods were used including purposive,\textsuperscript{30, 31, 33} convenience,\textsuperscript{35, 39} random,\textsuperscript{32, 40, 41} and a priori.\textsuperscript{28} Although nine papers did not explicitly state their sampling method, it appeared from the results in four cases that a purposive/convenience sample was used\textsuperscript{26, 31, 37, 42} and in one that an a-priori sample was employed.\textsuperscript{36} Survey response rates ranged from 37% to 77%.\textsuperscript{38} with a median response of 58.5%.
Participant characteristics
The majority of participants were medical doctors working in either primary or secondary care.26–28 30–36 40–43 Other participants had a background in nursing,26 30 public health,38 39 or pharmacy.35 40 Some studies only acknowledged the use of ‘stakeholders’ who were drawn from relevant organizations, and another did not specify the background of their participants other than stating that they were healthcare providers.29

Data shared/linkage
The extent of proposed data sharing/linkage differed widely between studies. The majority sought to identify views to data sharing for research,27 29–32 34 36 41 43 two on pilots of data sharing26 35 and the remainder on hypothetical scenarios.28 33 37–40 42

The most common linkages were between laboratory and radiology records,26–28 30 32 34 36 followed by patient records from either primary or secondary care.27 31 34 37 41 Other studies described shared medication data which included data from computerized physician order entry systems, discharge summaries, prescriptions, and medication lists.27 28 30–32 35 37 40 42 Three studies did not specify the clinical data to be shared.29 38 41

Type and purpose of data sharing/linkage
One study addressed views to sharing data with patients,26 and there was also sharing of data at an aggregate level, that is, anonymised data, with public health agencies.29 33 38 39 The majority of papers described data sharing (hypothetical and real) at the individual patient level, usually between pharmacies, and primary and secondary care, and in the USA between pharmacies and insurance companies.27–32 34–37 40–43

Views on data sharing/linkage
Although the majority of views expressed about data sharing were positive, there were some negative views.28–32 34 35 38–40 43 Studies categorized as ‘undecided’ in online supplementary table S1 reported both positive and negative views which appeared to be influenced by prior experience, as those with experience of linked data were generally in favor and those without were generally negative.27 33 41 A full range of views about data sharing were not identified in two studies as their purpose was to identify potential barriers to data sharing.36 37

The use of patient data for public health purposes was described by four papers: Rudin et al29 described the views of clinicians about sharing their data with public health departments, AbdelMalik et al38 discussed the need for patient-identifiable data for public health and the restrictions imposed by current legislation, and El Emam et al33 and Heidebrecht et al39 discussed the use of data for assessing immunization coverage. These studies showed that HCPs were positive about data sharing for a public health purpose. Only one study found that the view towards the secondary use of patient data would depend on the degree of identifiability.33

Barriers to data sharing/linkage
The key findings of each study are summarized in online supplementary table S1. A frequently mentioned barrier to data sharing related to start-up and maintenance costs, including remuneration for participating providers.26 27 31 32 34 38 Concerns about data governance were also common, including data security, legal restrictions, and data quality.27–29 31–33 38 40 Technical problems such as lack of interoperability between IT systems were also identified,29 31 37 40 42 although Paré et al37 reported that these were less of an issue for data sharing per se. Privacy issues were cited in four studies.31 32 38 42 Consent was seen as necessary, although this was deemed impractical to obtain for large anonymised whole population studies in one study,33 and as a potential barrier by a minority (19%) of participants in another.34

Figure 1 PRISMA flowchart of literature review (based on Moher et al25).
As a group, physicians often suggested possible interference with their patient–physician relationships, threats to their professional autonomy, and a potential increase in the use of such data for litigation. Participants also reported a potential for an increased workload associated with uploading, verifying, and updating data. Lack of awareness of Health Information Exchange, that is, the use of electronic movement of health-related information among organizations according to nationally (US) recognized standards, was identified as a barrier to data sharing, as was being a non-user of existing infrastructures, such as electronic prescribing systems. The lack of a shared vision or commitment from management and competition between healthcare providers were also cited as obstacles.

### Facilitators to data sharing/linkage

Several studies listed possible improvements in patient care and safety as facilitators for data sharing. Rudin et al. identified trust in the system as both a barrier and a facilitator as concerns of physicians about the sharing of clinical data appeared to be less in those who used linked IT based patient information systems. El Emam et al. identified several governance features, such as comprehensive data sharing agreements, the use of de-identified data, and mandatory reporting, for example public health purposes that would reassure HCPs and facilitate support for research with linked data.

The involvement of the relevant HCPs in the development of data sharing procedures would appear to facilitate data sharing, as would perceived ownership of any given project. Being a current user of a data sharing system, having a preference to view health records electronically, involvement in quality reporting initiatives, or perceived improvements in patient care associated with data sharing all acted as facilitators.

HCPs perceived data sharing as beneficial, and with the potential of reducing healthcare costs by saving clinician time in accessing relevant patient data and providing timely access to comprehensive whole population trends and longitudinal data. Fontaine et al. found that using a precursor system, that is, introducing a ‘light’ version of the planned system before full engagement, helped to dispel concerns of HCPs. Other facilitators included the clinical usefulness of the system, a well designed and easy to use interface, reliable system performance, and the ability to give and receive feedback.

### DISCUSSION

#### Summary of evidence

Data sharing was generally supported, and particularly so if HCPs had prior experience of its application as explicitly described in one study which compared the differences in views between those HCPs with and without experience in data sharing. Although no study identified a solely negative stance towards the sharing or linking of clinical data, several barriers were identified. Set-up costs for the required hardware and internet links along with subsequent system maintenance were perceived to be a problem, despite the potential for a reduction in healthcare costs overall.

Potential improvements in patient care and safety were seen as facilitators of data sharing in contrast to lack of any perceived usefulness and patient benefits acting as deterrents for participation. Hence, system utility and performance would be keys for a successful data sharing project as they can act as both barriers and facilitators. The results from this review indicate that HCPs are unlikely to support any data linkage system that is complicated, time-consuming, or costly. On the other hand, if benefits could be demonstrated, for example by providing easy access to comprehensive and longitudinal data, and particularly if the data were able to support strategic goals, this would work as a facilitator for data sharing. The direct involvement of HCP ‘champions’ willing to drive the project, in particular from doctors, nurses, and pharmacists, was identified to be crucial for success, as were willingness to co-operate, involvement, and psychological ownership which led to more enthusiasm.

### Limitations

This review summarizes a heterogeneous set of studies from different countries, with different methodologies and different data sharing or linkage schemes. Thus, only limited generalization and interpretation of the data is possible. Although the quality of included papers was not formally assessed using standard scales, any uncertainties about the exact nature of the study were found, authors were contacted to clarify information about study methods, such as number of interviews conducted, nature of data sharing, and clarification of participants’ characteristics.

Duplicate data extraction was not performed. However, a random sample of both included and excluded papers (n=32) was discussed. In addition, to facilitate confidence in reliability, a selection of papers was assessed twice by the same researcher, several weeks apart, and outcomes in terms of paper inclusion or data extracted were identical.

Data sharing was associated with different terms, particularly in the USA, including electronic medical records, electronic health records, health information exchange, and computer information systems, often without explanation for readers unfamiliar with the described setting. One of the terms used to describe data sharing was ‘computerized physician/provider order entry’ (CPOE), a term that was excluded as closer examination of the papers using this term showed that CPOE described solely the sharing of information about a patient within a single setting, generally between the wards and pharmacy of a single hospital. The search strategy was not amended in order to include new terms identified during the review, such as ‘electronic health record’ (EHR), as they were already being successfully identified by the original search strategy. However, this might have led to a failure to identify all possible relevant papers.

It was also not always clear whether electronic health records (EHRs) were shared within a single setting, such as a single practice or hospital or across settings, that is, record accessibility in primary and secondary care. This level of detail was central as the aim of the current review was to identify views of HCPs on data sharing and linkage across health sectors to inform the development of the CHIMES program for developing a more efficient system for pharmacovigilance. Although authors were contacted if information within their published paper was inadequate for an inclusion/exclusion decision, not every author provided the requested information, which led to five papers being excluded.

Other problems encountered were the inclusion of views from mixed populations of HCPs, managers, and the lay public and in which answers provided could not be attributed by population group. This limited the validity of the results, as reported views might have been voiced by a healthcare manager or system administrator without a clinical background.
CONCLUSION
Most studies described the sharing of data at an individual patient level and hence the observed views about barriers and facilitators to data sharing concerned the sharing of patient identifiable data. Identified barriers included costs, governance issues, and a perceived interference with the prescriber/patient relationship. Facilitators to data sharing were direct involvement of relevant HCPs in system design and the accessibility, perceived usefulness, and potential perceived benefits of the system. Benefits included easy access to complete and comprehensive patient data and the potential for improving quality of care and patient safety. In general, the views of HCPs were positive towards data sharing and linkage but the identified barriers will have to be addressed in future data linkage projects to facilitate support of healthcare professionals.

Contributors PJH was the chief investigator of the overall program, conceived the research and led writing of the proposal for funding. CB and JH were co-investigators and led the writing of the work package which included this literature review. YMH was responsible for the draft of the literature review protocol, the daily study conduct and co-ordination, acquisition of data, analysis, producing tables and figures and interpretation of data. YMH drafted/co-led writing of the paper and incorporating feedback from co-authors on successive drafts. CB and JF contributed to the literature review protocol design and subsequent analysis, and co-led the writing of the paper. All authors commented on the initial drafts of the paper and revision of successive drafts. The final version of the manuscript was approved by all authors.

Funding This work was supported by the Chief Scientist Office (Child Medical Records for Safer Medicines (CHIMES) Applied Research Program, grant number ARPG/07/4).

Competing interests: None.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/.

REFERENCES