BMJ Open

Core requirements for successful data linkage – an example of a triangulation method

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011879
Article Type:	Research
Date Submitted by the Author:	11-Mar-2016
Complete List of Authors:	Hopf, Yvonne; Klinikum der Universitat Munchen, Pharmacy; University of Aberdeen, Centre of Academic Primary Care Francis, Jill; City University London, Health Services Research and Management Division Helms, Peter; University of Aberdeen, Child Health Haughney, John; University of Aberdeen, Academic Primary Care Bond, Christine; University of Aberdeen, ;
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Health services research, Qualitative research
Keywords:	mixed methods research, data linkage, triangulation

Title: Core requirements for successful data linkage – an example of a triangulation method

Type of submission: Research article

Authors: Hopf Y M[1], Francis J[2], Helms P J[3], Haughney J[4], Bond C[5]

[1] Yvonne Marina Hopf, Centre of Academic Primary Care, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK, yhopf@med.lmu.de

[2] Jill Francis, School of Health Sciences, City University London, Northampton Square, London, EC1V 0HB, UK, Jill.Francis.1@city.ac.uk

[3] Peter J. Helms, Child Health, University of Aberdeen, Royal Aberdeen Children's Hospital, Westburn Road, Aberdeen AB25 2ZG, UK, p.j.helms@abdn.ac.uk

[4] John Haughney, Centre of Academic Primary Care, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK, j.haughney@abdn.ac.uk

[5] Christine Bond, Centre of Academic Primary Care, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK, c.m.bond@abdn.ac.uk

Corresponding Author (present address): Yvonne Marina Hopf, Klinikum der Universität München, Apotheke, Pettenkofer Str 8a, 80336 München, yhopf@med.lmu.de , Telephone: 0049 89 4400 543761, Fax: 0049 89 4400 53716

Keywords: mixed methods research; data linkage; triangulation

Word Count (main text only): 3586

Number of tables: 2

Number of figures: 1

ABSTRACT

Objectives

The aim was to explore the views of professional stakeholders and healthcare professionals (HCPs) on the linkage of UK National Health Service (NHS) data for paediatric pharmacovigilance purposes and to make recommendations for such a system.

Methods

A mixed methods approach including a literature review, focus-groups and a three-round Delphi survey with HCPs in Scotland was followed by a triangulation process using a systematic, validated protocol. The survey was structured using the Theoretical Domains Framework of behaviour change. Items retained after applying the matrix-based triangulation process were thematically coded. Ethical approval was granted by the North of Scotland Research Ethics Service.

Results

A total of 25 key findings from all four studies were identified during triangulation. There was good consensus; 21 key findings were agreed and remained to inform recommendations. The items were coded as practical/technical (*e.g.* decision about the unique patient identifier to use), mandatory (*e.g.* governed by statute), essential (consistently mentioned in all studies and therefore needed to ensure professional support) or preferable.

Conclusion

The development of a paediatric linked database has support from professional stakeholders and HCPs in Scotland. The triangulation identified three sets of core requirements for a new system of data linkage. An additionally fourth set of 'preferable' requirements might increase engagement of HCPs and their support for the new system.

Strengths and limitations of this study

- The mixed methods approach of this study utilised several methods to employ the strengths while addressing the inherent weaknesses of each method.
- No validated instruments were used; the interview and focus group schedules were based on the initial literature review.
- The use of a theoretical approach in the form of the Theoretical Domains Framework added robustness to the results of the Delphi, allowed for a systematic exploration of potential problems
- The use of a triangulation protocol allowed for a systematic comparison of findings between the different methods used.

BACKGROUND & SIGNIFICANCE

Adverse drug reactions (ADRs) are defined by the World Health Organisation[1] as noxious and unintended responses to a drug which occur at doses normally used in man. ADRs are more likely to occur when medicines are used off-label[2]. Off-label use of medication refers to all uses of a marketed drug not detailed in the Summary of Product Characteristics including therapeutic indication, use in age-subsets, appropriate strength (dosage), pharmaceutical form and route of administration. This is of particular concern in children who receive proportionately more off-label and unlicensed medicines[2-4] than other patient groups. Reported figures differ, but it is estimated that five out of eight severe ADRs in paediatric inpatients are related to off-label use of drugs in children[5].

Reports of ADRs are collected nationally in many countries *e.g.* Yellow Card Scheme in the UK or the FDA Adverse Event Reporting System database (FAERS) in the US. However it is widely acknowledged that these systems underestimate the true prevalence of ADRs[6, 7]. A systematic review of papers published between 1986 and 2006 on the views of healthcare professionals (HCPs) towards ADR reporting found that non-reporting was related to ignorance and lack of knowledge of the reporting system (95%), lack of time or similar reasons (77%), fear of filing an inappropriate report (72%), indifference and uncertainty about causality (67%), and the perception that licensed drugs are safe (47%[7]). New approaches of signal generation to enhance existing systems are required.

There is increasing electronic collection and storage of healthcare data during routine clinical practice for the purposes of more efficient clinical communication and health service administration. However, this contemporary and comprehensive information resource could also facilitate new methodological approaches in population health research, especially pharmacovigilance. Small scale studies have already shown that identification of suspected ADRs is possible using this routine data[8].

It is unclear, however, how stakeholders such as patients or HCPs would view the secondary use of routinely collected healthcare data, especially if individual datasets were linked.

OBJECTIVE

The aim of this study was explore the views and opinions of HCPS on linking routinely collected data and make evidence based recommendations for a system of pharmacovigilance based on linked data. The study was part of the CHIMES (Child Medical Records for Safer Medicines) programme in Scotland. The overall aim of the programme was to develop a novel paediatric pharmacovigilance system based on linkage of routinely collected data in primary (GP and community pharmacy) and secondary (hospital) care.

METHODS & MATERIALS

This was a mixed methods study which included data from: a literature review[9], an interview study[10], a focus group study[11] and a Delphi survey[12]. All these studies have been previously reported. In this paper the results of the four stages of data collection are formally triangulated, and recommendations for a new system of pharmacovigilance are made. All studies were approved by the North of Scotland Research Ethics Service and NHS Research and Development.

Summaries of the methods used in the previously reported papers are presented below.

The systematic literature review of published papers described the views of healthcare professionals to data sharing and linkage. Searches were performed in Medline, EMBASE, SCOPUS, CINAHL and PsychINFO for papers in English between 2001 and 2011[9].

The interviews explored the views of a purposive sample of Scottish stakeholders on pharmacovigilance, confidentiality/patient privacy, data protection, acceptable and non-acceptable usage of data and dissemination of findings. Interviews were semi-structured, were audio-recorded and fully transcribed. A framework approach was used to identify themes inductively, but the analysis approach also allowed for the identification of emergent themes[10].

Focus groups were conducted with pharmacists, nurses, general practitioners and paediatricians. Stratified purposive sampling was used to ensure inclusion of participants from a range of urban, rural, and remote settings and geographic locations across Scotland. Invitation packs were mailed to potential participants. Building on the interviews, the focus groups explored views of the proposed linkage and identified perceived barriers. All focus groups were audio recorded and transcribed verbatim. Themes were identified via a framework approach[11].

Based on the findings of the two qualitative studies a three-round Delphi survey with HCPs (pharmacists, nurses and medical doctors with an interest in paediatric medicine) in Scotland was conducted to identify consensus on essential system components required for HCPs to support or facilitate the proposed data linkage and continue recording the necessary data[12]. The survey was structured using the Theoretical Domains Framework[13] of behaviour change. A systematic data reduction exercise based on methods proposed by Prior *et al.*[14] was applied between Round 1 and 2. Consensus criteria for each round were set *a priori* at 66.7% for Round 2 and 90% for Round 3.

A triangulation protocol based on the methods proposed by Farmer *et al.*[15] was applied to interpret and integrate key findings from the literature review and the three empirical studies in order to identify essential system components. These methods proceeded in three steps. First, a matrix was constructed to allow comparison of key issues against as represented in the individual studies. Second, themes were compared to create a single list of issues, barriers and facilitators based on the individual study results. Third, agreement between studies was coded using a convergence coding scheme as presented in Table 1. 'Agreement' indicates that the key finding was identified in a particular study, 'partial agreement' means that the finding was partially covered, 'disagreement' indicates a contradictory finding. If none of these three codes could be attributed, the label 'silence' was used.

Findings were additionally categorised as either practical/technical, mandatory, essential, or preferable. Mandatory items were those underpinned by regulation or statute/legislation. Essential items were those in which there was (at least partial) consensus across at least three of the four studies (as listed in Table 2) and preferred items were those that did not match the above definitions but were deemed to increase the stakeholders' support and active engagement with the system.

RESULTS

In the literature review a total of 2917 titles were screened and 18 papers were included, describing the views of HCPs on linking or sharing healthcare data at an individual patient level[9]. These showed that data sharing at a patient identifiable level was more common than population based data linkage. Data sharing was often described within specific settings, such as an emergency department or a specialist ward or the pharmacy department of a hospital. HCPs who reported a previous positive experience with such data sharing perceived a positive impact on their work and patient safety and were more likely to support data sharing. However, funding (start-up and maintenance), technical problems (compatible IT systems) and governance issues were identified as potential barriers to successful implementation of data sharing.

Between February and October 2010, 25 participants were interviewed from 11 of the 24 Scottish Health Boards[10]. Interviewees had positive views on the proposal to use routinely collected data to create a pharmacovigilance resource for children in Scotland. Practical, ethical, and legal issues were identified related to the sharing of the data, the ownership of the linked data, consent for data linkage, anonymisation, and confidentiality. These results were further explored in focus groups with health practitioners.

Between August 2010 and May 2011, six focus groups were conducted with 22 participants from seven different Scottish Health Boards[11]. Participants were recruited from all professional backgrounds and from both primary and secondary care. Half of the participants were pharmacists

(n=11, 50%) and the majority were female (n=13, 59%). Focus group participants reported that governance problems and relevant legislation should be addressed. They proposed that patients would have issues with the use of their data for research purposes, with key areas of sensitivity including confidentiality and third party access to patient identifiable data. Participants supported the proposed linkage, and their range of views was similar to those of the national interviewees. However, more concerns were voiced about funding of such data linkage, particularly if this meant any diversion of funds from their own area of practice.

The first round of the Delphi study generated over 1000 individual statements from 61 participants. After systematic item reduction 149 items were retained for the second round, in which participants were asked to rate their level of agreement with each item. Items reaching the consensus criterion of 66.7% were entered into the third round. After the third round, the retained consensus items focused on professional standards, requirements for linkage and the use and form of potential feedback. Overall the results confirmed that participants were generally willing to facilitate the proposed linkage, dependent on adherence to professional standards, relevant legislation, ethical approval, secure data sharing agreements, and support from their employers.

A total of 25 key findings were identified across all four studies and the inter-study level of agreement is shown in Table 2. Findings showed near perfect agreement (full or partial) between the qualitative studies (23/25, 92%). No disagreement on any finding was noted between the qualitative studies, but each of the qualitative studies had one coding of 'silence': Focus group participants did not discuss whether data ownership would have to be clarified, and interview participants did not identify feedback from linkage studies as a facilitator of future support. Four of the 25 findings from the qualitative studies were not carried forward in the Delphi survey, *i.e.* the pre-defined consensus criteria were not reached; of the remainder, three results (3/25, 12%) from the survey findings did not reflect the findings from the qualitative arm, namely the role of the UK National Health Service

(NHS) as a facilitator of the data linkage, whether benefits outweighed the risks, or if professional guidelines would prevent data sharing.

All but four of the findings identified in the empirical studies were also identified in the systematic review[9]. Understanding of pharmacovigilance and understanding of the current system were not identified in the literature review, nor were the two Scottish-specific findings (NHS as a facilitator or the use of the Community Health Index (CHI) as a unique patient identifier).

Figure 1 indicates how the findings resulting from the triangulation process are mapped out displaying the conceptual model for the new system implementation.

DISCUSSION

Main findings

The triangulation process identified three core sets of requirements necessary for the implementation of a new linked data system for paediatric pharmacovigilance, including 'practical/technical' requirements as well as 'mandatory' requirements such as ethical approval, and 'essential' requirements such as transparency. A fourth set of requirements, labelled 'preferable', might increase the engagement of HCPs with such a new system.

Strengths & limitations

Triangulation has been described since the 1970s[16] and was initially defined as "*the combination of methodologies in the study of the same phenomenon*" (Denzin, 1978, p. 291, as seen in [16, 17]). The combination can be across different data sources, investigators, theories, or methodologies data collection methods[16, 17]. O'Cathain *et al.*[18] defined triangulation as investigating the same issue using different methods - an approach well suited to mixed methods research. The use of a triangulation protocol allowed for a systematic comparison of findings between the different methods used. The use of a mixed methods approach allowed minority issues in the interviews to be carried forward, *i.e.* issues that were only mentioned by one or two participants. The Delphi survey results that were

based on those findings confirmed overlapping findings between the different studies, providing face validity and robustness of the findings against method-based variation. The initial Delphi questionnaire was developed based on the findings from the interview and focus group study in order to present those findings to a larger audience with the aim of gaining consensus on the most important items. Although the questionnaire was not psychometrically validated it incorporated the use of a tested and validated framework allowing for a systematic assessment of barriers to data linkage, the Theoretical Domains Framework (TDF)[13]. The TDF accounted for all barriers identified in the qualitative studies, hence confirming good empirical coverage. Utilising several methods gave the capacity to employ the strengths and address the inherent weaknesses of each method.

This paper has not considered the opinions of patients/the public towards the secondary use of routinely collected data; these have been explored in a parallel study which is currently being analysed.

Views towards routinely linked NHS data

The proposed data linkage was perceived to be useful for pharmacovigilance research across all methods used in this study. Small scale studies have already indicated that it is possible to use linked data for epidemiology research[19-22] and that linked data can be used successfully for signal generation relevant to possible ADRs[8]. Participants indicated conditional support for the proposed linkage, with conditions ranging from anonymisation of the data to ethical approval as well as adherence to legislation and professional codes of conduct. A range of potentially relevant legal frameworks was identified during the interviews including the Data Protection Act, the Common Law Duty of Confidentiality, and the Human Rights Act[10]. In the qualitative studies there was confusion as to which legislation, which guidelines and standards were directly relevant to the proposed data linkage[10]. Several statements included potential regulatory and legislative frameworks but did not discuss any of these in detail as it was assumed that 'relevant' laws and guidelines would be followed. Despite probing, participants in focus groups did not provide further information on the

term 'relevant,' referring to experts that would be better suited to answer these questions despite the fact that even those experts interviewed commented that these had rarely been formally tested in a court of law.

The definition of 'relevant' is not only important for the application of legislation and guidelines but also for the anonymisation of data. Participants also used the term 'adequate' in the context of anonymisation of data security measures. However, neither the reviewed literature nor the empirical work provided a working definition of these terms or criteria for 'adequacy'. Not only would it be necessary to have an agreed definition in order to apply 'relevant' and 'adequate' measures, but it would also be important to specify who should decide what 'relevant' and 'adequate' means.

The findings suggest that the benefits of the proposed data linkage would outweigh the risks to patient privacy and confidentiality although the concerns about possible risks might have caused participants in all three methods to conclude that "*Ethical approval is required*". However participants did not specify what kind of ethical approval would be expected. The options would be approval by the National Research Ethics Service (NRES) and/or an internal ethics committee specific for the linked data. Some of the data providers in Scotland, such as the Health Information Centre (HIC) in Tayside or the NHS Information Service Division (ISD), already have internal ethics approval boards that assess requests for data access[23]. The internal approval board for ISD, the Privacy Advisory Committee (PAC), additionally demands that any requests to access patient identifiable data or data linkage projects are reviewed by an NHS ethics committee prior to submission to PAC[24]. In conjunction with the expressed views of HCPs that the NHS should be a facilitator of the data linkage, this could mean that the standards employed by NHS ISD as the biggest data collector of the Scottish NHS would find widespread HCP approval. However, this and the other two suggestions of an independent user like PAC or the government, *e.g.* the Department of Health, was not endorsed in the Delphi survey. No consensus emerged indicating a high level of uncertainty amongst the HCP

community about the governance of linked health data. The findings from this work indicate an awareness that a certain amount of patient identifiable data would be necessary in order for the proposed data linkage system to function. This finding is not new, as a Wellcome Trust report[25] has highlighted that the use of clinical data is rarely fully anonymous. Current standards employed by the NHS do recognize this, as approval is required from NHS ethics committees, and requests for patient identifiable data will have to be approved by relevant Caldicott guardians to ensure that only the minimum amount of identifiable data is used to answer the questions being posed[26]. A Caldicott guardian is the person responsible within NHS structures in the UK to ensure adherence to a set of principles and processes which provide a framework of quality standards for the management of confidentiality and access to patient information[26]. The General Medical Council (UK) would prefer anonymised disclosures of data although acknowledges that the release of patient identifiable information might be justified in the public interest[27].

Requirements for data linkage

Based on these findings the following might be appropriate to consider before implementing a system of pharmacovigilance based on linked health care data:

1) practical/technical: storage space for data, software compatibility
2) mandatory: 'to comply with law' and professional guidance
3) essential: 'to ensure support of HCPs'
4) preferable: 'to increase support and active engagement with the system'

'Practical' requirements include the need for meaningful data, in order to have data that is useful and the provision of server space for storage and processing the linked data. A Wellcome Trust report published in 2009 suggested the use of 'safe havens' for data linkage projects[25]. Software compatibility across and within the different Health Boards in Scotland was also questioned by

participants. The lack of a single standard system for reporting and recording medical data has been noted previously[28]. General agreement was that any technical or practical problems could be solved.

Compliance with relevant legislation was seen as 'mandatory'. The following acts were identified as relevant for the planned project: the Data Protection Act 1998 and the Common Law Duty of Confidentiality. In addition to the law compliance with professional standards and ethical approval were deemed to be important. For all recommendations, clear definitions are required. This would be in addition to the terms 'relevant' and 'adequate' as discussed above.

'Essential' requirements include system design options that would ensure the support of HCPs for the planned data linkage. One of these requirements would be that the planned data linkage should not impact negatively on the current workload, which is in line with previous findings[29-33]. Clear information governance structures were requested across all arms of this study. Concerns about governance including data security, legal restrictions and data quality are not new and have been cited previously as potential barriers to data sharing[31, 34-40]. Potential governance structures could relate to consent and confidentiality structures as well as the use of data sharing agreements. Similar features have been recently described by El Emam *et al.*[40] as facilitators for data sharing. Participants also requested transparency about the kind of data used and the research initiated and informing the public about the proposed system.

'Preferable' items would increase the support from HCPs, hopefully leading to a more active engagement in the new system design. Feedback from the linked data resources was seen as a strong motivator for HCPs. Frontline HCPs saw feedback as an opportunity to review their own practice. Feedback was requested in an 'easy-to-digest' format with the possibility of accessing further information if required. In particular interview participants saw feedback as a chance to enable data providers to increase data quality and with this the quality of the linked data. In this context Aylin *et al.*[20] found that the quality of the data sent to the Hospital Episodes Statistics in

England improved over time as a consequence of regular feedback to participating hospitals. The qualitative studies identified the belief that if the feedback provided ultimately benefitted patients, HCPs would become more interested in ensuring high quality data entry. In terms of consent, participants saw a discussion about consent as necessary, although they concluded that requesting explicit consent from all patients in Scotland would be unfeasible. This is in contrast to Patel's findings[33] where only a minority of participants thought the request of individual consent would be too complicated and time-consuming. An opt-out possibility was deemed acceptable for patients, although HCPs requested that this assumption should be confirmed with the general public prior to the implementation of the new system. Finally, in particular frontline HCPs would prefer to have clear data sharing agreements in place which is in line with previous findings of El Emam *et al.*[40].

Taking all four categories into account, the first three would describe the basis for the proposed new system, with the fourth increasing support from HCPs (as detailed in Figure 1).

CONCLUSION

The proposed data linkage was perceived as addressing a gap in current knowledge. Nonetheless, participants did identify a range of problems and concerns that should be addressed prior to implementation of such a system. Adherence to relevant legislation, such as the Data Protection Act, and the application of appropriate governance, such as adherence to professional standards and the Caldicott principles, were deemed to be important in ensuring confidentiality for patients and prescribers. Although no consensus was reached in the first two studies on the need for consent, the Delphi study identified that the offer of an 'opt-out option' for patients would have support from HCPs.

Three key areas (labelled practical/technical, mandatory and essential)were identified that would require attention prior to system introduction although the terms 'relevant' and 'appropriate' would need to be better defined in order to ensure a strong and secure basis for the proposed new system

for pharmacovigilance in children. A fourth category (labelled preferable) described conditions that may help to increase support of HCPs for such a new system.

The work presented in this paper demonstrates that the development of a paediatric linked database for pharmacovigilance has support from professional stakeholders and HCPs in Scotland together with specific recommendations to inform the design of such a system.

Acknowledgements:

We would like to thank all our participants for taking part in the studies.

Declaration of conflicting interests:
There are no conflicting interests.

Funding Statement

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

CONTRIBUTORSHIP STATEMENT:

PH was the Chief Investigator of the overall programme (CHIMES), 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 the literature review and the empirical studies. YH was responsible for the draft of the overall study protocol as well as the triangulation protocol, the daily study conduct and co-ordination, acquisition of data, analysis, producing tables and figures and interpretation of data. YH drafted/co-led writing of the paper and incorporating feedback from co-authors on successive drafts. CB and JF contributed to study 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.

DATA SHARING

Extra data such as the search strategy of the literature review, the study information sheets and consent forms for all studies, as well as the Delphi questionnaires or the interview and focus group topic guides, example of descriptive chart (interviews/focus groups), and examples for categorisation of statements in Delphi are available by emailing Yvonne Hopf (corresponding author).

REFERENCES

1. World Health Organization. *Safety of Medicines-A guide to detecting and reporting adverse drug reactions*. Geneva: Word Health Organiszation, 2002.
2. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. *Br J Clin Pharmacol* 2001;52:77-83.
3. McIntyre J, Conroy S, Avery A, Corns H, Choonara I. Unlicensed and off label prescribing of drugs in general practice. *Arch Dis Child* 2000;83(6):498-501. PubMed PMID: 11087285. Pubmed Central PMCID: 1718565.
4. Ekins-Daukes S, Helms PJ, Simpson CR, Taylor MW, McLay JS. Off-label prescribing to children in primary care: retrospective observational study. *Eur J Clin Pharmacol* 2004 Jul;60(5):349-53. PubMed PMID: 15146302.
5. Choonara I, Conroy S. Unlicensed and off-label drug use in children - Implications for safety. *Drug Saf* 2002;25(1):1-5. PubMed PMID: WOS:000174312500001.
6. BMA Board of Science. *Reporting adverse drug reactions - A guide for healthcare professionals*. London: British Medical Association, 2006.
7. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. *Drug Saf* 2009;32(1):19-31. PubMed PMID: 19132802.
8. Sun AP, Kirby B, Black C, Helms PJ, Bennie M, McLay JS. Unplanned medication discontinuation as a potential pharmacovigilance signal: a nested young person cohort study. *BMC Pharmacol Toxicol* 2014;15(11):11. PubMed PMID: 24594374. Pubmed Central PMCID: 3946605.
9. Hopf YM, Bond C, Francis J, Haughney J, Helms PJ. Views of healthcare professionals to linkage of routinely collected healthcare data: a systematic literature review. *J Am Med Inform Assoc* 2014;21(e1):e6-10. PubMed PMID: 23715802. Pubmed Central PMCID: PMC3957379. Epub 2013/05/30.
10. Hopf YM, Bond CB, Francis JJ, Haughney J, Helms PJ. Linked health data for pharmacovigilance in children: perceived legal and ethical issues for stakeholders and data guardians. *BMJ Open* 2014;4(2).
11. Hopf YM, Bond C, Francis J, Haughney J, Helms PJ. 'The more you link, the more you risk …' – a focus group study exploring views about data linkage for pharmacovigilance. *Br J Clin Pharmacol* 2014;78(5):1143-50.
12. Hopf YM, Francis J, Bond C, Haughney J, Helms PJ. Linking NHS data for pediatric pharmacovigilance: Results of a Delphi survey. *Res Soc Admin Pharm* 2016;12:267-280.

13. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A, et al. Making psychological theory useful for implementing evidence based practice: a consensus approach. *Qual Saf Health Care* 2005;14(1):26-33. PubMed PMID: 15692000. Pubmed Central PMCID: 1743963.
14. Prior ME, Hamzah JC, Francis JJ, Ramsay CR, Castillo MM, Campbell SE, et al. Pre-validation methods for developing a patient reported outcome instrument. *BMC Med Res Methodol* 2011;11:112. PubMed PMID: 21827689. Pubmed Central PMCID: 3225127.
15. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. *Qual Health Res* 2006;16(3):377-94. PubMed PMID: 16449687.
16. Burke Johnson R, Onwuegbuzie AJ, L.A. T. Towards a Definition of Mixed Methods Research. *J Mixed Method Res* 2007;1:112-32.
17. Silverman D. *Interpreting Qualitative Data*. 3rd ed. London: SAGE publication; 2006.
18. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *Br Med J* 2010;341:c4587. PubMed PMID: 20851841.
19. Helms PJ, Ekins Daukes S, Taylor MW, Simpson CR, McLay JS. Utility of routinely acquired primary care data for paediatric disease epidemiology and pharmacoepidemiology. *Br J Clin Pharmacol* 2005;59(6):684-90. PubMed PMID: 15948933. Pubmed Central PMCID: 1884863.
20. Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *Br Med J* 2007;334.
21. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf* 2007;16(4):393-401. PubMed PMID: 17066486.
22. Wettermark B, Hammar N, MichaelFored C, Leimanis A, Otterblad Olausson P, Bergman U, *et al.* The new Swedish Prescribed Drug Register—Opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf* 2007;16(7):726-35.
23. ISD Scotland. Information Services Division (ISD). Available at: www.isdscotland.org (accessed last 2 March 2015).
24. Privacy Advisory Committee. *Guidance Notes (for PAC applications)*. 2009.
25. Wellcome Trust. *Towards Consensus for Best Practice: Use of patient records from general practice for research*. London: Wellcome Trust, 2009.
26. Department of Health. *The Caldicott Guardian Manual 2010*. Leeds: Department of Health (DH/Digital Information Policy), 2010.
27. General Medical Council. *Confidentialty*. London: General Medical Council, 2009. http:// http://www.gmc-uk.org/guidance/ethical_guidance/confidentiality.asp (accessed 2 March 2015).
28. de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Fam Pract* 2006;23(2):253-63.
29. Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin CT. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: Patient and physician experiences. *J Am Med Inform Assoc* 2004;11(5):410-7.
30. Jaatinen P, Aarnio P, Asikainen P. The foundations for a regional information system based on a reference register. *J Inform Technol Healthcare* 2006;4(3):154-64.
31. Shapiro JS, Kannry J, Kushniruk AW, Kuperman G, New York Clinical Information Exchange Clinical Advisory S. Emergency physicians' perceptions of health information exchange. *J Am Med Inform Assoc* 2007;14(6):700-5. PubMed PMID: 17712079. Pubmed Central PMCID: 2213478.
32. Heidebrecht C, Foisy J, Pereira J, Quan S, Willison D, Deeks S, et al. Perceptions of immunization information systems for collecting pandemic H1N1 immunization data within

Canada's public health community: A qualitative study. *BMC Public Health*. 2010;10(1):523. PubMed PMID: doi:10.1186/1471-2458-10-523.

33. Patel V, Abramson EL, Edwards A, Malhotra S, Kaushal R. Physicians' potential use and preferences related to health information exchange. *Int J Med Inf* 2011;80(3):171-80. PubMed PMID: 21156351

34. Porteous T, Bond C, Robertson R, Hannaford P, Reiter E. *Electronic transfer of prescription-related information: comparing views of patients, general practitioners, and pharmacists*. *BR J Gen Prac* 2003;53:204-9.

35. Grossman JM, Bodenheimer TS, McKenzie K. Hospital-physician portals: the role of competition in driving clinical data exchange. *Health Aff (Millwood)* 2006;25(6):1629-36. PubMed PMID: 17102188.

36. AbdelMalik P, Boulos MN, Jones R. The perceived impact of location privacy: a web-based survey of public health perspectives and requirements in the UK and Canada. *BMC Public Health* 2008;8:156. PubMed PMID: 18471295. Pubmed Central PMCID: 2396622.

37. Rudin RS, Simon SR, Volk LA, Tripathi M, Bates D. Understanding the decisions and values of stakeholders in health information exchanges: experiences from Massachusetts. *Am J Public Health* 2009;99(5):950-5. PubMed PMID: 19299671. Pubmed Central PMCID: 2667860.

38. Fontaine P, Zink T, Boyle RG, Kralewski J. Health information exchange: participation by Minnesota primary care practices. *Arch Int Med* 2010;170(7):622-9. PubMed PMID: 20386006.

39. Wright A, Soran C, Jenter CA, Volk LA, Bates DW, Simon SR. Physician attitudes toward health information exchange: results of a statewide survey. *J Am Med Inform Assoc* 2010;17(1):66-70. PubMed PMID: 20064804. Pubmed Central PMCID: 2995629.

40. El Emam K, Mercer J, Moreau K, Grava-Gubins I, Buckeridge D, Jonker E. Physician privacy concerns when disclosing patient data for public health purposes during a pandemic influenza outbreak. *BMC Public Health* 2011;11:454. PubMed PMID: 21658256. Pubmed Central PMCID: 3130674.

Tables

Coding label	Convergence coding
Agreement	Finding has been identified
Partial Agreement	Finding is covered partially
Disagreement	Finding is contradicted
Silence	Finding does not appear

Table 1: Convergence coding scheme for triangulation protocol (based on Farmer *et al.* 2006).

Table 2: Overview of identified key issues and how they triangulate across the literature review and the three studies ('Agreement = key finding has been identified; 'Partial agreement = finding is partially covered; 'disagreement = contradictory statement; 'Silence' = not apparent; 'not addressed' =subject not addressed in study; Bold font indicates agreement across three or more arms). *NB*: This table extends to 2 pages.

Key Finding	Literature Review	Interviews	Focus groups	Delphi Survey
Limited knowledge of available databases.	**Agreement**	**Agreement**	**Agreement**	**Agreement**
Understanding of pharmacovigilance based on WHO definition.	Not addressed	**Partial Agreement**	**Partial Agreement**	Not addressed
Limitations of current pharmacovigilance systems.	Not addressed	**Agreement**	**Agreement**	Not addressed
Data linkage is useful for pharmacovigilance research.	**Agreement**	**Agreement**	**Partial Agreement**	**Partial Agreement**
Adherence to legislation is necessary.	**Agreement**	**Agreement**	**Agreement**	Not addressed
It is not clear which legislation is relevant.	**Partial Agreement**	**Agreement**	**Agreement**	Not addressed
Data quality is important for data linkage.	**Agreement**	**Agreement**	**Agreement**	Partial **Consensus**
The responsibility for the linked data lies with the data owner.	Silence	**Agreement**	**Partial Agreement**	Silence
Data ownership needs to be clarified.	**Partial Agreement**	**Agreement**	Silence	**Agreement**
The NHS should be a facilitator of the data linkage.	Silence	**Partial Agreement**	**Agreement**	Disagreement
The benefits of the data linkage outweigh the risks.	**Partial Agreement**	**Agreement**	**Partial Agreement**	Disagreement [1]
Professional guidelines might prevent data sharing.	Silence	**Agreement**	**Partial Agreement**	Disagreement
Conditional support if data linkage available.	**Agreement**	**Agreement**	**Agreement**	**Agreement**
Information governance is a facilitator of data linkage.	**Agreement**	**Agreement**	**Agreement**	**Agreement**

[1] disagreement only in the way that this was mentioned but dropped across the rounds, indicating that there was no consensus on this item, hence disagreement rather than silence

Key Finding	Literature Review	Interviews	Focus groups	Delphi Survey
Feedback from studies using the linked data is a facilitator.	Agreement	Silence	Agreement	Agreement
Anonymisation is mandatory.	Partial Agreement	Partial Agreement	Agreement	Agreement
Information of and input from the public is important.	Agreement	Agreement	Agreement	Agreement
An opt-out option for patients is acceptable.	Agreement	Partial Agreement	Partial Agreement	Agreement
Technical problems can be solved.	Agreement	Agreement	Agreement	Agreement
Use of CHI number is acceptable.	Silence	Agreement	Agreement	Agreement
Safeguards like sanctions should be used.	Agreement	Agreement	Partial Agreement	Silence
Ethical approval will be required.	Agreement	Agreement	Agreement	Agreement
Data linkage should not impact negatively on current workload.	Agreement	Agreement	Agreement	Agreement
A certain amount of patient identifiable data will be necessary for the linkage.	Agreement	Agreement	Agreement	Silence
Assurance of confidentiality would facilitate support.	Agreement	Partial Agreement	Partial Agreement	Agreement

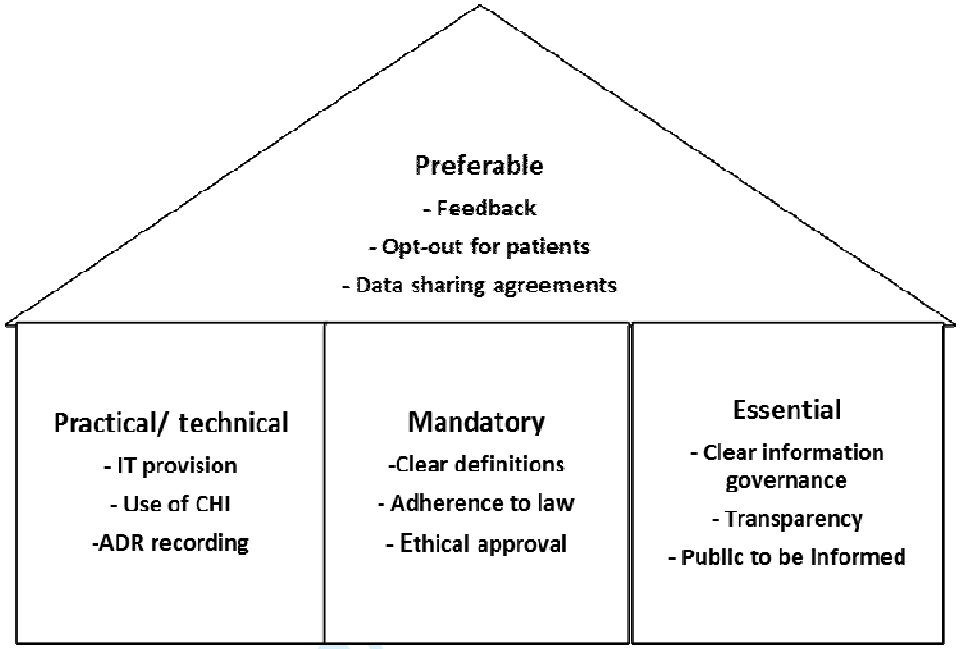


Figure 1: Building blocks of recommendations for the proposed data linkage system. Practical requirements describe practical and technical necessities for the actual data linkage process; mandatory requirements describe compliance with relevant legislation; essential requirements describe system design options; and preferable items describe items that would increase the support from HCPs. CHI-Community Health Index, ADR-Adverse Drug Reaction.

BMJ Open

Core requirements for successful data linkage – an example of a triangulation method

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011879.R1
Article Type:	Research
Date Submitted by the Author:	24-Aug-2016
Complete List of Authors:	Hopf, Yvonne; Klinikum der Universitat Munchen, Pharmacy; University of Aberdeen, Centre of Academic Primary Care Francis, Jill; City University London, Health Services Research and Management Division Helms, Peter; University of Aberdeen, Child Health Haughney, John; University of Aberdeen, Academic Primary Care Bond, Christine; University of Aberdeen, ;
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Health services research, Qualitative research
Keywords:	mixed methods research, data linkage, triangulation

Title: Core requirements for successful data linkage – an example of a triangulation method

Type of submission: Research article

Authors: Hopf Y M[1], Francis J[2], Helms P J[3], Haughney J[4], Bond C[5]

[1] Yvonne Marina Hopf, Centre of Academic Primary Care, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK, yhopf@med.lmu.de

[2] Jill Francis, School of Health Sciences, City University London, Northampton Square, London, EC1V 0HB, UK, Jill.Francis.1@city.ac.uk

[3] Peter J. Helms, Child Health, University of Aberdeen, Royal Aberdeen Children's Hospital, Westburn Road, Aberdeen AB25 2ZG, UK, p.j.helms@abdn.ac.uk

[4] John Haughney, Centre of Academic Primary Care, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK, j.haughney@abdn.ac.uk

[5] Christine Bond, Centre of Academic Primary Care, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK, c.m.bond@abdn.ac.uk

Corresponding Author (present address): Yvonne Marina Hopf, Klinikum der Universität München, Apotheke, Pettenkofer Str 8a, 80336 München, yhopf@med.lmu.de , Telephone: 0049 89 4400 543761, Fax: 0049 89 4400 53716

Keywords: mixed methods research; data linkage; triangulation

Word Count (main text only): 4060

Number of tables: 2

Number of figures: 1

ABSTRACT

Objectives

The aim was to explore the views of professional stakeholders and healthcare professionals (HCPs) on the linkage of UK National Health Service (NHS) data for paediatric pharmacovigilance purposes and to make recommendations for such a system.

Methods

A mixed methods approach including a literature review, interviews, focus-groups and a three-round Delphi survey with HCPs in Scotland was followed by a triangulation process using a systematic protocol. The survey was structured using the Theoretical Domains Framework of behaviour change. Items retained after applying the matrix-based triangulation process were thematically coded. Ethical approval was granted by the North of Scotland Research Ethics Service.

Results

Results from 18 papers, 23 interviewees, 23 participants of focus groups and 61 completed questionnaires in the Delphi survey contributed to the triangulation process. A total of 25 key findings from all four studies were identified during triangulation. There was good convergence; 21 key findings were agreed and remained to inform recommendations. The items were coded as practical/technical (*e.g.* decision about the unique patient identifier to use), mandatory (*e.g.* governed by statute), essential (consistently mentioned in all studies and therefore needed to ensure professional support) or preferable.

Conclusion

The development of a paediatric linked database has support from professional stakeholders and HCPs in Scotland. The triangulation identified three sets of core requirements for a new system of data linkage. An additionally fourth set of 'preferable' requirements might increase engagement of HCPs and their support for the new system.

Strengths and limitations of this study

- The mixed methods approach of this study utilised several methods to employ the strengths while addressing the inherent weaknesses of each method.
- No validated instruments were used; the interview and focus group schedules were based on the initial literature review.
- The use of a theoretical approach in the form of the Theoretical Domains Framework added robustness to the results of the Delphi, allowed for a systematic exploration of potential problems
- The use of a triangulation protocol allowed for a systematic comparison of findings between the different methods used.

BACKGROUND & SIGNIFICANCE

Adverse drug reactions (ADRs) are defined by the World Health Organisation[1] as noxious and unintended responses to a drug which occur at doses normally used in man. ADRs are more likely to occur when medicines are used off-label[2]. Off-label use of medication refers to all uses of a marketed drug not detailed in the Summary of Product Characteristics (the legally binding product information in the UK for medication) including therapeutic indication, use in age-subsets, appropriate strength (dosage), pharmaceutical form and route of administration. This is of particular concern in children, who receive proportionately more off-label and unlicensed medicines[2-4] than other patient groups. Reported figures differ, but it is estimated that five out of eight severe ADRs in paediatric inpatients are related to off-label use of drugs in children[5].

Reports of ADRs are collected nationally in many countries *e.g.* Yellow Card Scheme in the United Kingdom (UK) or the Food & Drug Administration (FDA) Adverse Event Reporting System database (FAERS) in the United States (US). However it is widely acknowledged that these systems underestimate the true prevalence of ADRs[6, 7]. A systematic review of papers published between 1986 and 2006 on the views of healthcare professionals (HCPs) towards ADR reporting found that non-reporting was related to ignorance and lack of knowledge of the reporting system (95%), lack of time or similar reasons (77%), fear of filing an inappropriate report (72%), indifference and uncertainty about causality (67%), and the perception that licensed drugs are safe (47%[7]). New approaches of signal generation to enhance existing systems are required.

There is increasing electronic collection and storage of healthcare data during routine clinical practice for the purposes of more efficient clinical communication and health service administration. However, this contemporary and comprehensive information resource could also facilitate new methodological approaches in population health research, especially pharmacovigilance. Small scale studies have already shown that identification of suspected ADRs is possible using this routine data[8].

It is unclear, however, how stakeholders such as national (Scottish) experts on pharmacovigilance, confidentiality or data protection and HCPs would view the secondary use of routinely collected healthcare data, especially if individual datasets were linked. . The research reported here was part of the CHIMES (Child Medical Records for Safer Medicines) programme in Scotland. The overall aim of this programme was to develop a novel paediatric pharmacovigilance system based on linkage of routinely collected data in primary (GP and community pharmacy) and secondary (hospital) care.to The aim of this research was to provide recommendations for such a national resource that would allow monitoring for long term outcomes of exposure to medicines during childhood and adolescence. Adverse reactions to medicines could be identified by stop dates, reduction of initial doses or switches of medication. In line with use of other existing datasets derived from routinely collected data, our assumption was that access to the data by researchers would only be allowed after review of the application by an internal ethics, or equivalent, committee.

OBJECTIVE

The aim of this study was to explore the views and opinions of HCPs on linking routinely collected data and make evidence based recommendations for a system of pharmacovigilance based on linked data.

METHODS & MATERIALS

This was a mixed methods study which included data from: a literature review[9], an interview study[10], a focus group study[11] and a Delphi survey[12]. All these studies have been previously reported. In this paper the results of the four stages of data collection are formally triangulated, and recommendations for a new system of pharmacovigilance are made. All studies were approved by the North of Scotland Research Ethics Service and NHS Research and Development.

Summaries of the methods used in the previously reported papers are presented below.

The systematic literature review of published papers described the views of healthcare professionals to data sharing and linkage. Searches were performed in Medline, EMBASE, SCOPUS, CINAHL and PsychINFO for papers in English between 2001 and 2011[9].

The interviews explored the views of a purposive sample of Scottish stakeholders on pharmacovigilance, confidentiality/patient privacy, data protection, acceptable and non-acceptable usage of data and dissemination of findings. Interviews were semi-structured, were audio-recorded and fully transcribed. A framework approach was used to identify themes inductively, but the analysis approach also allowed for the identification of emergent themes[10].

Focus groups were conducted with pharmacists, nurses, general practitioners and paediatricians, from both, primary and secondary care. Stratified purposive sampling was used to ensure inclusion of participants from a range of urban, rural, and remote settings and geographic locations across Scotland. Paediatric experience was preferred but not essential. Invitation packs were mailed to potential participants. Building on the interviews, the focus groups explored views of the proposed linkage and identified perceived barriers. All focus groups were audio recorded and transcribed verbatim. Themes were identified via a framework approach[11].

Based on the findings of the two qualitative studies a three-round Delphi survey with HCPs (pharmacists, nurses and medical doctors with an interest in paediatric medicine) in Scotland was conducted to identify consensus on essential system components required for HCPs to support or facilitate the proposed data linkage and continue recording the necessary data[12]. A random sample was drawn from national sampling frames of the target populations, utilising for example the NHS Information Service Division workforce lists for general practitioners and practice nurses, a list of registered premises by the Practitioner Service Division for community pharmacists, or paediatricians via the Scottish Paediatric Society. The survey was structured using the Theoretical Domains Framework[13] of behaviour change. A systematic data reduction exercise based on methods

proposed by Prior *et al.*[14] was applied between Round 1 and 2. Consensus criteria for each round were set *a priori* at 66.7% for Round 2 and 90% for Round 3.

A triangulation protocol based on the methods proposed by Farmer *et al.*[15] was applied to interpret and integrate key findings from the literature review and the three empirical studies in order to identify essential system components. These methods proceeded in three steps. First, a matrix was constructed to allow comparison of key issues against those as represented in the individual studies. Second, themes were compared to create a single list of issues, barriers and facilitators based on the individual study results. Third, agreement between studies was coded using a convergence coding scheme as presented in Table 1. 'Agreement' indicates that the key finding was identified in a particular study, 'partial agreement' means that the finding was partially covered, 'disagreement' indicates a contradictory finding. If none of these three codes could be attributed, the label 'silence' was used.

RESULTS

In the literature review, which was conducted in 2011, a total of 2917 titles were screened and 18 papers were included, describing the views of HCPs on linking or sharing healthcare data at an individual patient level[9]. These showed that data sharing at a patient identifiable level was more common than population based data linkage. Data sharing was often described within specific settings, such as an emergency department or a specialist ward or the pharmacy department of a hospital. HCPs who reported a previous positive experience with such data sharing perceived a positive impact on their work and patient safety and were more likely to support data sharing. However, funding (start-up and maintenance), technical problems (compatible Information Technology -IT- systems) and governance issues were identified as potential barriers to successful implementation of data sharing.

Between February and October 2010, 25 participants were interviewed from 11 of the 24 Scottish Health Boards[10]. Interviewees had positive views on the proposal to use routinely collected data to create a pharmacovigilance resource for children in Scotland. Practical, ethical, and legal issues were identified related to the sharing of the data, the ownership of the linked data, consent for data linkage, anonymisation, and confidentiality. These results were further explored in focus groups with health practitioners.

Between August 2010 and May 2011, six focus groups were conducted with 22 participants from seven different Scottish Health Boards[11]. Participants were recruited from all professional backgrounds and from both primary and secondary care. Half of the participants were pharmacists (n=11, 50%) and the majority were female (n=13, 59%). Focus group participants reported that governance problems and relevant legislation should be addressed. They proposed that patients would have issues with the use of their data for research purposes, with key areas of sensitivity including confidentiality and third party access to patient identifiable data. Participants supported the proposed linkage, and their range of views was similar to those of the national interviewees. However, more concerns were voiced about funding of such data linkage, particularly if this meant any diversion of funds from their own area of practice.

The Delphi study was conducted in three rounds from August 2011 to February 2012[12]. The first round of the Delphi study generated over 1000 individual statements from 61 participants. After systematic item reduction 149 items were retained for the second round, in which participants were asked to rate their level of agreement with each item. Items reaching the consensus criterion of 66.7% were entered into the third round. After the third round, the retained consensus items focused on professional standards, requirements for linkage and the use and form of potential feedback. Overall the results confirmed that participants were generally willing to facilitate the proposed linkage, dependent on adherence to professional standards, relevant legislation, ethical approval, secure data sharing agreements, and support from their employers.

Applying the triangulation protocol, a total of 25 key findings were identified across all four studies and the inter-study level of agreement is shown in Table 2. Findings showed near perfect agreement (full or partial) between the qualitative studies (23/25, 92%). No disagreement on any finding was noted between the qualitative studies, but each of the qualitative studies had one coding of 'silence': Focus group participants did not discuss whether data ownership would have to be clarified, and interview participants did not identify feedback from linkage studies as a facilitator of future support. Four of the 25 findings (16%) from the qualitative studies were not carried forward in the Delphi survey, *i.e.* the pre-defined consensus criteria were not reached; of the remainder, three results (3/25, 12%) from the survey findings did not reflect the findings from the qualitative arm, namely the role of the UK National Health Service (NHS) as a facilitator of the data linkage, whether benefits outweighed the risks, or if professional guidelines would prevent data sharing.

All but four of the findings identified in the empirical studies were also identified in the systematic review[9]. Understanding of pharmacovigilance and understanding of the current system were not identified in the literature review, nor were the two Scotland-specific findings (NHS as a facilitator or the use of the Community Health Index (CHI) as a unique patient identifier).

Essential system components were deduced from the Delphi analysis plus key findings that showed (partial) agreement across at least three studies as listed in Table 1. Statutory items, such as legislation were included as following the law would be the basis of any decision. Key findings were sorted into inductively and resulted in four categories: practical/technical, mandatory, essential, or preferable. Practical requirements include the need for meaningful data, the provision of server space for storage and processing the linked data as well as software compatibility across and within the different Health Boards in Scotland. General agreement was that any technical or practical problems could be solved. Mandatory items were those underpinned by regulation or statute/legislation. Literature and participants identified that following 'relevant' legislation is important but a definition of relevant was seldom provided or discussed. Focus group participants

declined requests to be more specific and referred to legal specialists instead. Indeed, the interview partners with legal expertise identified the widest range of legislative acts and non-statutory guidelines that should be observed. Additionally, participants deemed ethical approval to be important. Essential items were those in which there was (at least partial) consensus across at least three of the four studies (as listed in Table 2), *e.g.* that the planned data linkage should not impact negatively on the current workload or clear information governance structures. Participants stipulated that transparency included informing the public about the new system. Items coded as 'preferable' were those that did not match the above definitions but were deemed to increase the stakeholders' support and active engagement with the system. Clear data sharing agreements in place as well as feedback from the linked data resources and an opt-out possibility for patients was deemed acceptable, but HCPS requested to evaluate patients' opinions prior to the implementation of the new system. Figure 1 indicates how the findings resulting from the triangulation process are mapped out displaying the conceptual model for the new system implementation.

DISCUSSION

Main findings

The triangulation process identified three core sets of requirements necessary for the implementation of a new linked data system for paediatric pharmacovigilance, including 'practical/technical' requirements as well as 'mandatory' requirements such as ethical approval, and 'essential' requirements such as transparency. A fourth set of requirements, labelled 'preferable', might increase the engagement of HCPs with such a new system.

Strengths & limitations

Triangulation has been described since the 1970s[16] and was initially defined as "*the combination of methodologies in the study of the same phenomenon*" (Denzin, 1978, p. 291, as seen in [16, 17]). The combination can be across different data sources, investigators, theories, methodologies, or data

collection methods[16, 17]. O'Cathain *et al.*[18] defined triangulation as investigating the same issue using different methods - an approach well suited to mixed methods research. The use of a triangulation protocol allowed for a systematic comparison of findings between the different methods used. The use of a mixed methods approach allowed minority issues in the interviews to be carried forward, *i.e.* issues that were only mentioned by one or two participants. The Delphi survey results that were based on those findings confirmed overlapping findings between the different studies, providing face validity and robustness of the findings against method-based variation. The initial Delphi questionnaire was developed based on the findings from the interview and focus group study in order to present those findings to a larger audience with the aim of gaining consensus on the most important items. Although the questionnaire was not psychometrically validated it incorporated the use of a tested and validated framework allowing for a systematic assessment of barriers to data linkage, the Theoretical Domains Framework (TDF)[13]. The TDF accounted for all barriers identified in the qualitative studies, hence confirming good empirical coverage. Utilising several methods gave the capacity to employ the strengths and address the inherent weaknesses of each method.

The single studies involved in the triangulation took place over a period of 2 years. Some of the opinions from the stakeholders and HCPs might have changed over that time. We did not re-test their views and opinions a second time. We are however confident that this triangulation work does indeed present a good starting point of the requirements for such a suggested data linkage project. Once the proposed system has been implemented, surveys amongst data providers and data users should be performed on a regular basis to ensure that the way the data linkage is performed and analysed still acceptable for HCPs.

This paper has not considered the opinions of patients/the public towards the secondary use of routinely collected data; these have been explored in a parallel study which is currently being analysed.

Views towards routinely linked NHS data

The proposed data linkage was perceived to be useful for pharmacovigilance research across all methods used in this study. Small scale studies have already indicated that it is possible to use linked data for epidemiology research[19-22] and that linked data can be used successfully for signal generation relevant to possible ADRs[8]. Participants indicated conditional support for the proposed linkage, with conditions ranging from anonymisation of the data to ethical approval as well as adherence to legislation and professional codes of conduct. A range of potentially relevant legal frameworks was identified during the interviews including the Data Protection Act 1998, the Common Law Duty of Confidentiality, and the Human Rights Act 1998[10]. In the qualitative studies there was confusion as to which legislation, which guidelines and standards were directly relevant to the proposed data linkage[10]. Several statements included potential regulatory and legislative frameworks but did not discuss any of these in detail as it was assumed that 'relevant' laws and guidelines would be followed. Despite probing, participants in focus groups did not provide further information on the term 'relevant,' referring to experts that would be better suited to answer these questions despite the fact that even those experts interviewed commented that these had rarely been formally tested in a court of law.

The definition of 'relevant' is not only important for the application of legislation and guidelines but also for the anonymisation of data. Participants also used the term 'adequate' in the context of anonymisation of data security measures. However, neither the reviewed literature nor the empirical work provided a working definition of these terms or criteria for 'adequacy'. Not only would it be necessary to have an agreed definition in order to apply 'relevant' and 'adequate' measures, but it would also be important to specify who should decide what 'relevant' and 'adequate' means.

The findings suggest that the benefits of the proposed data linkage would outweigh the risks to patient privacy and confidentiality although the concerns about possible risks might have caused

participants in all three methods to conclude that "*Ethical approval is required*". However participants did not specify what kind of ethical approval would be expected. The options would be approval by the National Research Ethics Service (NRES) and/or an internal ethics committee specific for the linked data. Some of the data providers in Scotland, such as the Health Information Centre (HIC) in Tayside or the NHS Information Service Division (ISD), already have internal ethics approval boards that assess requests for data access[23]. The internal approval board for ISD, the Privacy Advisory Committee (PAC), additionally demands that any requests to access patient identifiable data or data linkage projects are reviewed by an NHS ethics committee prior to submission to PAC[24]. In conjunction with the expressed views of HCPs that the NHS should be a facilitator of the data linkage, this could mean that the standards employed by NHS ISD as the biggest data collector of the Scottish NHS would find widespread HCP approval. However, this and the other two suggestions of an independent user like PAC or the government, *e.g.* the Department of Health, was not endorsed in the Delphi survey. No consensus emerged indicating a high level of uncertainty amongst the HCP community about the governance of linked health data. The findings from this work indicate an awareness that a certain amount of patient identifiable data would be necessary in order for the proposed data linkage system to function. This finding is not new, as a Wellcome Trust report[25] has highlighted that the use of clinical data is rarely fully anonymous. Current standards employed by the NHS do recognize this, as approval is required from NHS ethics committees, and requests for patient identifiable data will have to be approved by relevant Caldicott guardians to ensure that only the minimum amount of identifiable data is used to answer the questions being posed[26]. A Caldicott guardian is the person responsible within NHS structures in the UK to ensure adherence to a set of principles and processes which provide a framework of quality standards for the management of confidentiality and access to patient information[26]. The General Medical Council (UK) would prefer anonymised disclosures of data although acknowledges that the release of patient identifiable information might be justified in the public interest[27].

Requirements for data linkage

Based on these findings the following might be appropriate to consider before implementing a system of pharmacovigilance based on linked health care data:

1) practical/technical: storage space for data, software compatibility
2) mandatory: to comply with law and professional guidance
3) essential: to ensure support of HCPs
4) preferable: to increase support and active engagement with the system

'Practical' requirements include the need for meaningful data, in order to have data that is useful and the provision of server space for storage and processing the linked data. A Wellcome Trust report published in 2009 suggested the use of 'safe havens' for data linkage projects[25]. Software compatibility across and within the different Health Boards in Scotland was also questioned by participants. The lack of a single standard system for reporting and recording medical data has been noted previously[28]. Despite this, results indicated that any technical or practical problems could be resolved.

Compliance with relevant legislation was seen as 'mandatory'. The following acts were identified as relevant for the planned project: the Data Protection Act 1998 and the Common Law Duty of Confidentiality. In addition to the law compliance with professional standards and ethical approval were deemed to be important. For all recommendations, clear definitions are required. This would be in addition to the terms 'relevant' and 'adequate' as discussed above.

'Essential' requirements describe system options that would ensure the support of HCPs for the planned data linkage. One of these requirements would be that the planned data linkage should not impact negatively on the current workload, which is in line with previous findings[29-33]. Clear information governance structures were requested across all arms of this study. Concerns about governance including data security, legal restrictions and data quality are not new and have been

cited previously as potential barriers to data sharing[31, 34-40]. Potential governance structures could relate to consent and confidentiality structures as well as the use of data sharing agreements. Similar features have been recently described by El Emam *et al.*[40] as facilitators for data sharing. Participants also requested transparency about the kind of data used and the research initiated and informing the public about the proposed system.

'Preferable' items would increase the support from HCPs, hopefully leading to a more active engagement in the new system design. Feedback from the linked data resources was seen as a strong motivator for HCPs. Frontline HCPs saw feedback as an opportunity to review their own practice. Feedback was requested in an 'easy-to-digest' format with the possibility of accessing further information if required. In particular interview participants saw feedback as a chance to enable data providers to increase data quality and with this the quality of the linked data. In this context Aylin *et al.*[20] found that the quality of the data sent to the Hospital Episodes Statistics in England improved over time as a consequence of regular feedback to participating hospitals. The qualitative studies identified the belief that if the feedback provided ultimately benefitted patients, HCPs would become more interested in ensuring high quality data entry. In terms of consent, participants saw a discussion about consent as necessary, although they concluded that requesting explicit consent from all patients in Scotland would be unfeasible. This is in contrast to Patel's findings[33] where only a minority of participants thought the request of individual consent would be too complicated and time-consuming. An opt-out possibility was deemed acceptable for patients, although HCPs requested that this assumption should be confirmed with the general public prior to the implementation of the new system. Finally, in particular frontline HCPs would prefer to have clear data sharing agreements in place which is in line with previous findings of El Emam *et al.*[40].

Taking all four categories into account, the first three would describe the basis for the proposed new system, with the fourth increasing support from HCPs (as detailed in Figure 1).

CONCLUSION

The proposed data linkage was perceived as addressing a gap in current knowledge. Nonetheless, participants did identify a range of problems and concerns that should be addressed prior to implementation of such a system. Adherence to relevant legislation, such as the Data Protection Act, and the application of appropriate governance, such as adherence to professional standards and the Caldicott principles, were deemed to be important in ensuring confidentiality for patients and prescribers. Although no consensus was reached in the first two studies on the need for consent, the Delphi study identified that the offer of an 'opt-out option' for patients would have support from HCPs.

The work presented in this paper demonstrates that the development of a paediatric linked database for pharmacovigilance has support from professional stakeholders and HCPs in Scotland together with specific recommendations to inform the design of such a system.

Acknowledgements:

We would like to thank all our participants for taking part in the studies.

Declaration of conflicting interests:

There are no conflicting interests.

Funding Statement

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

CONTRIBUTORSHIP STATEMENT:

PH was the Chief Investigator of the overall programme (CHIMES), 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 the literature review and the empirical studies. YH was responsible for the

draft of the overall study protocol as well as the triangulation protocol, the daily study conduct and co-ordination, acquisition of data, analysis, producing tables and figures and interpretation of data. YH drafted/co-led writing of the paper and incorporating feedback from co-authors on successive drafts. CB and JF contributed to study 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.

Data Sharing Statement

Extra data such as the search strategy of the literature review, the study information sheets and consent forms for all studies, as well as the Delphi questionnaires or the interview and focus group topic guides, example of descriptive chart (interviews/focus groups), and examples for categorisation of statements in Delphi are available by emailing Yvonne Hopf (corresponding author).

REFERENCES

1. World Health Organization. *Safety of Medicines-A guide to detecting and reporting adverse drug reactions*. Geneva: Word Health Organiszation, 2002
2. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. *Br J Clin Pharmacol* 2001;52:77-83.
3. McIntyre J, Conroy S, Avery A, Corns H, Choonara I. Unlicensed and off label prescribing of drugs in general practice. *Arch Dis Child* 2000;83(6):498-501. PubMed PMID: 11087285. Pubmed Central PMCID: 1718565.
4. Ekins-Daukes S, Helms PJ, Simpson CR, Taylor MW, McLay JS. Off-label prescribing to children in primary care: retrospective observational study. *Eur J Clin Pharmacol* 2004 Jul;60(5):349-53. PubMed PMID: 15146302.
5. Choonara I, Conroy S. Unlicensed and off-label drug use in children - Implications for safety. *Drug Saf* 2002;25(1):1-5. PubMed PMID: WOS:000174312500001.
6. BMA Board of Science. *Reporting adverse drug reactions - A guide for healthcare professionals*. London: British Medical Association, 2006.
7. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. *Drug Saf* 2009;32(1):19-31. PubMed PMID: 19132802.
8. Sun AP, Kirby B, Black C, Helms PJ, Bennie M, McLay JS. Unplanned medication discontinuation as a potential pharmacovigilance signal: a nested young person cohort study. *BMC Pharmacol Toxicol* 2014;15(11):11. PubMed PMID: 24594374. Pubmed Central PMCID: 3946605.
9. Hopf YM, Bond C, Francis J, Haughney J, Helms PJ. Views of healthcare professionals to linkage of routinely collected healthcare data: a systematic literature review. *J Am Med Inform Assoc* 2014;21(e1):e6-10. PubMed PMID: 23715802. Pubmed Central PMCID: PMC3957379. Epub 2013/05/30.
10. Hopf YM, Bond CB, Francis JJ, Haughney J, Helms PJ. Linked health data for pharmacovigilance in children: perceived legal and ethical issues for stakeholders and data guardians. *BMJ Open* 2014;4(2).
11. Hopf YM, Bond C, Francis J, Haughney J, Helms PJ. 'The more you link, the more you risk …' – a focus group study exploring views about data linkage for pharmacovigilance. *Br J Clin Pharmacol* 2014;78(5):1143-50.
12. Hopf YM, Francis J, Bond C, Haughney J, Helms PJ. Linking NHS data for pediatric pharmacovigilance: Results of a Delphi survey. *Res Soc Admin Pharm* 2016;12:267-280.
13. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A, et al. Making psychological theory useful for implementing evidence based practice: a consensus approach. *Qual Saf Health Care* 2005;14(1):26-33. PubMed PMID: 15692000. Pubmed Central PMCID: 1743963.
14. Prior ME, Hamzah JC, Francis JJ, Ramsay CR, Castillo MM, Campbell SE, et al. Pre-validation methods for developing a patient reported outcome instrument. *BMC Med Res Methodol* 2011;11:112. PubMed PMID: 21827689. Pubmed Central PMCID: 3225127.
15. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. *Qual Health Res* 2006;16(3):377-94. PubMed PMID: 16449687.
16. Burke Johnson R, Onwuegbuzie AJ, L.A. T. Towards a Definition of Mixed Methods Research. *J Mixed Method Res* 2007;1:112-32.
17. Silverman D. *Interpreting Qualitative Data.* 3rd ed. London: SAGE publication; 2006.
18. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *Br Med J* 2010;341:c4587. PubMed PMID: 20851841.

19. Helms PJ, Ekins Daukes S, Taylor MW, Simpson CR, McLay JS. Utility of routinely acquired primary care data for paediatric disease epidemiology and pharmacoepidemiology. *Br J Clin Pharmacol* 2005;59(6):684-90. PubMed PMID: 15948933. Pubmed Central PMCID: 1884863.
20. Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *Br Med J* 2007;334.
21. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf.* 2007;16:393-401. PubMed PMID: 17066486.
22. Wettermark B, Hammar N, MichaelFored C, Leimanis A, Otterblad Olausson P, Bergman U, *et al.* The new Swedish Prescribed Drug Register—Opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf* 2007;16(7):726-35.
23. ISD Scotland. Information Services Division (ISD). Available at: www.isdscotland.org (accessed last 2 March 2015).
24. Privacy Advisory Committee. *Guidance Notes (for PAC applications)*. 2009.
25. Wellcome Trust. *Towards Consensus for Best Practice: Use of patient records from general practice for research*. London: Wellcome Trust, 2009.
26. Department of Health. *The Caldicott Guardian Manual 2010*. Leeds: Department of Health (DH/Digital Information Policy), 2010.
27. General Medical Council. *Confidentialty*. London: General Medical Council, 2009. http:// http://www.gmc-uk.org/guidance/ethical_guidance/confidentiality.asp (accessed 2 March 2015).
28. de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Fam Pract* 2006;23:253-263.
29. Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin CT. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: Patient and physician experiences. *J Am Med Inform Assoc* 2004;11(5):410-7.
30. Jaatinen P, Aarnio P, Asikainen P. The foundations for a regional information system based on a reference register. *J Inform Technol Healthcare* 2006;4(3):154-64.
31. Shapiro JS, Kannry J, Kushniruk AW, Kuperman G, New York Clinical Information Exchange Clinical Advisory S. Emergency physicians' perceptions of health information exchange. *J Am Med Inform Assoc* 2007;14(6):700-5. PubMed PMID: 17712079. Pubmed Central PMCID: 2213478.
32. Heidebrecht C, Foisy J, Pereira J, Quan S, Willison D, Deeks S, et al. Perceptions of immunization information systems for collecting pandemic H1N1 immunization data within Canada's public health community: A qualitative study. *BMC Public Health*. 2010;10(1):523. PubMed PMID: doi:10.1186/1471-2458-10-523.
33. Patel V, Abramson EL, Edwards A, Malhotra S, Kaushal R. Physicians' potential use and preferences related to health information exchange. *Int J Med Inf* 2011;80(3):171-80. PubMed PMID: 21156351
34. Porteous T, Bond C, Robertson R, Hannaford P, Reiter E. Electronic transfer of prescription-related information: comparing views of patients, general practitioners, and pharmacists. *BR J Gen Prac* 2003;53:204-9.
35. Grossman JM, Bodenheimer TS, McKenzie K. Hospital-physician portals: the role of competition in driving clinical data exchange. *Health Aff (Millwood)* 2006;25(6):1629-36. PubMed PMID: 17102188.
36. AbdelMalik P, Boulos MN, Jones R. The perceived impact of location privacy: a web-based survey of public health perspectives and requirements in the UK and Canada. *BMC Public Health* 2008;8:156. PubMed PMID: 18471295. Pubmed Central PMCID: 2396622.

37. Rudin RS, Simon SR, Volk LA, Tripathi M, Bates D. Understanding the decisions and values of stakeholders in health information exchanges: experiences from Massachusetts. *Am J Public Health* 2009;99(5):950-5. PubMed PMID: 19299671. Pubmed Central PMCID: 2667860.

38. Fontaine P, Zink T, Boyle RG, Kralewski J. Health information exchange: participation by Minnesota primary care practices. *Arch Int Med* 2010;170(7):622-9. PubMed PMID: 20386006.

39. Wright A, Soran C, Jenter CA, Volk LA, Bates DW, Simon SR. Physician attitudes toward health information exchange: results of a statewide survey. *J Am Med Inform Assoc* 2010;17(1):66-70. PubMed PMID: 20064804. Pubmed Central PMCID: 2995629.

40. El Emam K, Mercer J, Moreau K, Grava-Gubins I, Buckeridge D, Jonker E. Physician privacy concerns when disclosing patient data for public health purposes during a pandemic influenza outbreak. *BMC Public Health* 2011;11:454. PubMed PMID: 21658256. Pubmed Central PMCID: 3130674.

Tables

Coding label	Convergence coding
Agreement	Finding has been identified
Partial Agreement	Finding is covered partially
Disagreement	Finding is contradicted
Silence	Finding does not appear

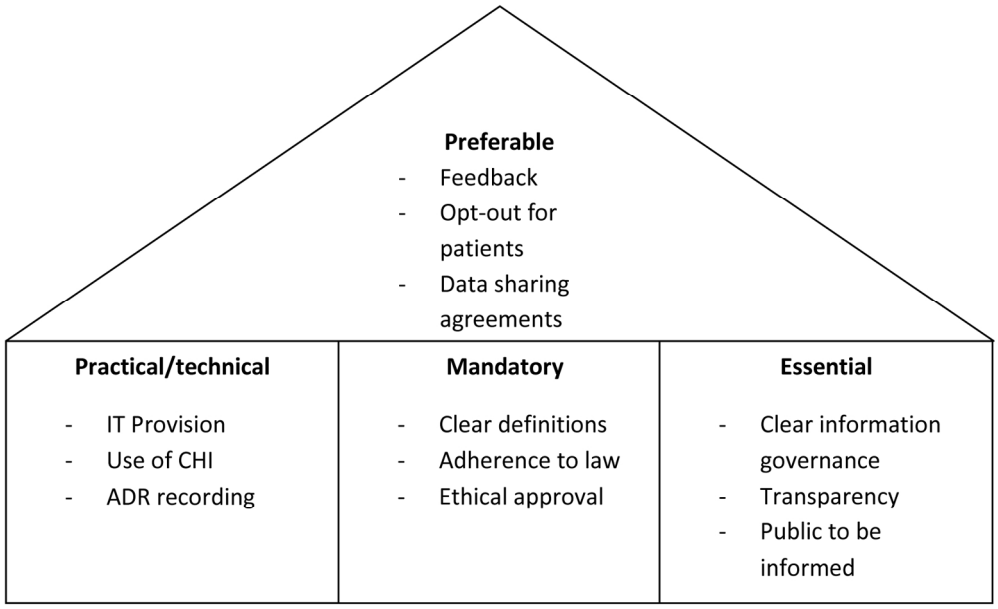
Table 1: Convergence coding scheme for triangulation protocol (based on Farmer *et al.* 2006).

Table 2: Overview of identified key issues and how they triangulate across the literature review and the three studies ('Agreement = key finding has been identified; 'Partial agreement = finding is partially covered; 'disagreement = contradictory statement; 'Silence' = not apparent; 'not addressed' =subject not addressed in study; Bold font indicates agreement across three or more arms). *NB*: This table extends to 2 pages.

Key Finding	Literature Review	Interviews	Focus groups	Delphi Survey
Limited knowledge of available databases.	**Agreement**	**Agreement**	**Agreement**	**Agreement**
Understanding of pharmacovigilance based on WHO definition.	Not addressed	Partial Agreement	Partial Agreement	Not addressed
Limitations of current pharmacovigilance systems.	Not addressed	Agreement	Agreement	Not addressed
Data linkage is useful for pharmacovigilance research.	**Agreement**	**Agreement**	**Partial Agreement**	**Partial Agreement**
Adherence to legislation is necessary.	**Agreement**	**Agreement**	**Agreement**	Not addressed
It is not clear which legislation is relevant.	**Partial Agreement**	**Agreement**	**Agreement**	Not addressed
Data quality is important for data linkage.	**Agreement**	**Agreement**	**Agreement**	Partial **Consensus**
The responsibility for the linked data lies with the data owner.	Silence	Agreement	Partial Agreement	Silence
Data ownership needs to be clarified.	**Partial Agreement**	**Agreement**	Silence	**Agreement**
The NHS should be a facilitator of the data linkage.	Silence	**Partial Agreement**	**Agreement**	Disagreement
The benefits of the data linkage outweigh the risks.	**Partial Agreement**	**Agreement**	**Partial Agreement**	Disagreement [1]
Professional guidelines might prevent data sharing.	Silence	Agreement	Partial Agreement	Disagreement
Conditional support if data linkage available.	**Agreement**	**Agreement**	**Agreement**	**Agreement**
Information governance is a facilitator of data linkage.	**Agreement**	**Agreement**	**Agreement**	**Agreement**

[1] disagreement only in the way that this was mentioned but dropped across the rounds, indicating that there was no consensus on this item, hence disagreement rather than silence

Key Finding	Literature Review	Interviews	Focus groups	Delphi Survey
Feedback from studies using the linked data is a facilitator.	Agreement	Silence	Agreement	Agreement
Anonymisation is mandatory.	Partial Agreement	Partial Agreement	Agreement	Agreement
Information of and input from the public is important.	Agreement	Agreement	Agreement	Agreement
An opt-out option for patients is acceptable.	Agreement	Partial Agreement	Partial Agreement	Agreement
Technical problems can be solved.	Agreement	Agreement	Agreement	Agreement
Use of CHI number is acceptable.	Silence	Agreement	Agreement	Agreement
Safeguards like sanctions should be used.	Agreement	Agreement	Partial Agreement	Silence
Ethical approval will be required.	Agreement	Agreement	Agreement	Agreement
Data linkage should not impact negatively on current workload.	Agreement	Agreement	Agreement	Agreement
A certain amount of patient identifiable data will be necessary for the linkage.	Agreement	Agreement	Agreement	Silence
Assurance of confidentiality would facilitate support.	Agreement	Partial Agreement	Partial Agreement	Agreement



Building blocks of recommendations for the proposed data linkage system. Practical requirements describe practical and technical necessities for the actual data linkage process; mandatory requirements describe compliance with relevant legislation; essential requirements describe system design options; and preferable items describe items that would increase the support from HCPs. ADR-Adverse Drug Reaction, CHI-Community Health Index, IT-Information Technology.

Figure 1

153x94mm (300 x 300 DPI)