

Title: Governance and oversight of researcher access to electronic health data: the role of the Independent Scientific Advisory Committee for MHRA database research (ISAC) 2006-2015

Authors:

Patrick Waller¹

Honorary Professor, Faculty of Epidemiology and Public Health, London School of Hygiene and Tropical Medicine

Jackie A Cassell²

Chair in Primary Care Epidemiology and Honorary Consultant in Public Health at Brighton and Sussex Medical School

Marcia Heinemann Saunders³

Chair, North West London Local Education and Training Board (Health Education England)

Richard Stevens⁴

Associate Professor, Medical Statistics Group, Nuffield Dept of Primary Care Health Sciences, University of Oxford

1. Chair, Independent Scientific Advisory Committee for MHRA database research (ISAC) from 2012 until January 2016
2. Scientific Member of ISAC from 2006, Acting Co-Chair 2010-11, Deputy Chair 2012-4
3. Lay Member of ISAC from 2006
4. Scientific Member of ISAC from 2009

Corresponding author:

Patrick Waller
153, Wolseley Road,
Rugeley,
Staffordshire WS15 2QT
United Kingdom
e-mail: patrick.waller@btinternet.com
Telephone: 01889 570417

Abstract

In order to promote understanding of UK governance and assurance relating to electronic health records (EHR) research, we here present and discuss the role of the Independent Scientific Advisory Committee for MHRA database research (ISAC) in evaluating protocols proposing the use of the Clinical Practice Research Datalink (CPRD). We describe the development of the Committee's activities during 2006-2015 alongside growth in data linkage and wider national EHR programmes, including the application and assessment processes, and our approach to undertaking this work. Our model can provide independence, challenge and support to data providers such as the CPRD database which has been used for well over a thousand medical research projects. ISAC's role in scientific oversight ensures feasible and scientifically acceptable plans are in place, while having both lay and professional membership addresses governance issues so as to protect the integrity of the database and ensure that public confidence is maintained.

Keywords

Electronic health records, Data security and confidentiality, Primary care, Databases and data mining, Record linkage

Introduction

The ISAC is a non-statutory government expert advisory body established in the UK in 2006 to provide advice to the Medicines and Healthcare Products Regulatory Agency (MHRA) on research related requests to access data provided from the General Practice Research Database (GPRD) and the Yellow Card Scheme for reporting of adverse drug reactions (ADRs), the main purpose of which is to identify signals of previously unrecognised ADRs. Professional members of the committee, such as clinicians and statisticians, make up the majority of the membership because our terms of reference emphasize scientific aspects of research in particular. However, the committee also includes two lay members, and is occasionally called upon to provide advice on broader issues such as health benefit, experience of UK primary care, and quality of patient information documentation. Prior to the formation of ISAC, requests to access GPRD for research purposes were subject to review by a Scientific and Ethical Advisory Group (SEAG) whose role was described briefly in 1998 [1].

In early 2012 GPRD was reconstituted as the Clinical Practice Research Datalink (CPRD), an observational data and interventional research service, jointly funded by the NHS National Institute for Health Research (NIHR) and the Medicines and Healthcare products Regulatory Agency (MHRA). The re-naming recognised the growing importance of data linkage beyond primary care, and the fact that CPRD services are designed to maximise the way anonymised NHS clinical data can be linked to enable many types of observational research and deliver research outputs that are beneficial to improving and safeguarding public health. The status of CPRD

was summarised by Herrett et al. in 2015 [2] who indicated that it has coverage of over 11.3 million patients from 674 practices with approximately 6.9% of the current UK population included. An important strength of the database is that these patients are broadly representative of the general population in terms of age, sex and ethnicity.

Overarching ethical approval has been granted for observational studies using anonymised CPRD data by a National Research Ethics Service committee, subject to review of each protocol by ISAC. Evaluation of protocols using Yellow Card data [3] follows similar principles to those described below but has formed a small and declining proportion of the work of the Committee (about 1% since 2012), and is not discussed further here.

As indicated on the ISAC section of the CPRD website [4], the purpose of the Committee's review of CPRD protocols is defined as "to ensure that investigators using the databases for research have feasible plans which do not raise governance concerns and reach an acceptable scientific standard. In this context we aim to provide timely, high quality peer review of protocols whilst recognising that the quality of the research ultimately remains the responsibility of the applicants."

Membership and development of ISAC

The Committee meets quarterly in the London offices of the MHRA but most of its work to review applications is performed between meetings, co-ordinated by its Secretariat and overseen by the chair. The size of the Committee has increased from 12 members in 2006 to 18 in 2015 (plus the chair). The Committee includes two lay members and 16 scientific members covering various disciplines including, in particular, epidemiology and biostatistics, along with a variety of clinical specialties e.g. paediatrics, diabetic medicine. The Committee publishes a summary of its minutes [5] and an annual report of its activities. Further information regarding the membership can be found in past annual reports which have been archived from the MHRA website [6]. Members are required to declare any potential conflicts of interest and, where these are significant, are disbarred from involvement in the review process. A modest fee and expenses are paid to members in relation to their attendance at meetings.

During the period 2006-11 all assessment was undertaken in the Committee with the Chair producing the feedback documents for all applications on the basis of member reviews. The MHRA provided secretarial staff resource only. The number of protocols increased from 114 in the first year to 155 during 2011 and, to address a further expected increase in volumes and complexity, plans for three changes were then formulated as follows: (1) a half-time paid chair role to assess and oversee the outcome of every application (2) a half-time scientific secretary who would undertake parallel reviews (3) a risk review system (see below) which would divide protocols into low risk (not requiring review by members but assessed by both ISAC chair and

scientific secretary) and medium or high risk which would also be subject to review by members. These changes were introduced in January 2012 and led to a reduction in mean review times from 30 working days in 2011 to 6 in 2014. During this period the number of new protocols submitted rose to 260 in 2015. The chair remained responsible for production of the feedback for all applications, including resubmissions and amendments, until July 2015 when a revised protocol review system was introduced (see below).

Application and assessment processes

All protocol submissions are accompanied by an application form which provides information on the applicants, their expertise and experience, funding source(s), conflicts of interest, publication plans and any request(s) for use of linked data. The form is periodically revised and the current version is available on the CPRD website [6]. Until mid-2015 applicants were encouraged by use of a checklist to structure their protocol according to a series of headings. Since then standard headings (see table 1) have become part of the application form so as to enable a simpler validation process which ensures that all the necessary sections have been included.

The application process is supported by detailed guidance which has developed considerably over time on the basis of experience and discussion at meetings. The current guidance can be found on the CPRD website [7].

The Committee's approach to the assessment of research protocols

Our overall approach is based on the tenet that observational research should be protocol driven and subject to defined standards as for other clinical research (e.g. trials). We endeavour to facilitate research by maintaining fast review times and avoiding backlogs, and we try to improve rather than prevent research which appears problematic, so few protocols are rejected outright. A system of "risk review" has been in place in recent years the purpose of which is to ensure that studies which are complex and have potentially bigger implications for public health receive a greater level of review than more basic studies. This is described in more detail below.

There is a risk that some research will be driven by the data and of selective publication of interesting findings derived by trawling through large amounts of data. We therefore try to limit and focus researchers, and generally avoid suggesting additional avenues of analysis unless these are aimed at exploring and improving validity (e.g. sensitivity analysis). Importantly, we try to encourage researchers to clearly specify in advance important aspects of the research such as case definitions and, above all, their approach to the analysis. Confounding is a potential concern in almost all observational research which is not purely descriptive and therefore warrants a section in an ISAC protocol in its own right. Most studies conducted in CPRD will be subject to some missing data and therefore we routinely require a plan to recognise and deal with it. The database itself has other limitations and we expect these to be recognised, along with specific consideration of limitations relating to the research questions being posed.

On the governance side, we make sure that applicants have sufficient expertise and experience in the use of primary care data, that conflicts of interest are declared and that any ethical concerns are addressed. The goal of maintaining full patient confidentiality is given very high priority and the potential risk of a breach is systematically assessed by ISAC for each protocol with mitigation being put in place where necessary (see below for details).

Risk review system

A system of "risk review" was implemented from January 2012 with the aim of rapidly categorising studies into those which do or do not require additional review by individual members of the Committee. Relatively straightforward protocols (e.g. basic descriptive epidemiology, studies of drug usage) have been designated low risk and those which are more complex designated medium or high risk if they meet one or more of the following criteria:

- Studies with potentially major public health/public interest implications
- Studies requiring contact with patients
- Studies with non-standard methods or analytical techniques
- Hypothesis testing drug safety studies

The risk designation is performed independently by both the scientific secretary and chair with the final decision being made by the chair. Judgement is involved in applying the criteria so as to ensure that those protocols which are most likely to benefit from additional review are those which receive it. In 2014, 74% of protocols submitted were judged low risk. The medium category risk category has rarely been applied and, since such protocols were being handled in the same way as those judged high risk, this category has now been dropped.

Further handling has depended on the risk categorisation, with low risk protocols being managed by the chair of ISAC and CPRD staff without reference to Committee

members. Protocols not deemed low risk are subjected to additional review by two members of the Committee selected by the chair on the basis of specific expertise, even work sharing as far as possible, and more detailed criteria for lay member and statistical member review. Members are allowed two weeks to complete their review and over the last 4 years, the response rate has exceeded 99%.

The scientific assessment from currently used for all applications is shown in Table 1. Scientific issues raised in the assessment are the major driver of the outcome and content of the feedback but a number of other considerations are also taken into account, as follows.

Ethical issues

A small proportion of protocols have raised significant ethical issues, usually because patients are to be approached for information or a sample. These have automatically been rated high risk and a lay member of the Committee is asked to review the protocol and any documentation intended for patients (e.g. consent forms, information sheets), and their comments included in the feedback. Such studies have also required review by research ethics committees.

The need for patient group involvement

Patient group involvement is encouraged by ISAC for most studies and is normally required for those which involve approaching patients directly and require specific ethical approval.

Funding source and conflict of interest statement

ISAC considers it important that the source of funding for a study is made explicit in the application and is consonant with a conflict of interest statement provided by the applicants.

Assessment of expertise and experience in the research team

Since mid-2015 all applicants have been required to provide a short curriculum vitae (CV) so that the expertise and experience of the research team can be assessed. We now also differentiate between the "chief investigator" who must be a senior researcher who takes overall responsibility for the research, and the "corresponding applicant" who submits the protocol and deals with the feedback. Unless it requires updating, a CV need only be submitted once following which it is given a reference number and allocated suffix(es) to indicate the person meets the criteria to be considered as one or more of the following:

C - Chief investigator (based on seniority)

E - Experienced CPRD user (has completed at least 3 studies)

P - UK primary care practitioner (e.g. GP, past or present)

S - Statistician (evidence of statistical training and experience)

L - experienced user of *linked* CPRD data

The combined expertise of the team is assessed in relation to the protocol in question and if the assessment of expertise and, if it appears deficient, an appropriate statement is included in the feedback e.g. "ISAC is concerned that the research team does not have sufficient experience in the use of primary care data and requires that a person with experience of practising primary care in the UK be added to the research team." Requests for access to the data are frequently submitted from abroad (notably from North America) and this issue is particularly important when none of the researchers are based in the UK, and they are likely to be unfamiliar with the healthcare system.

Issues related to the use of linked datasets

A subset of English practices (representing more than 50% all UK CPRD practices) have consented to participate in its linkage scheme. Patient-level data from consenting practices are linked via a trusted third party (the Health and Social Care Information Centre) to other existing data sources. Established linkages include inpatient and outpatient Hospital Episode Statistics, Office for National Statistics mortality data, and the Index of Multiple Deprivation and Townsend scores (data on socio-economic status). Bespoke linkages have included disease registries such as the National Cancer Registry and the Myocardial Ischaemia National Audit Project. Researchers can make requests for bespoke linkages for individual studies and further linkages are planned in the future.

Applicants are expected to explain why linked data are needed and how they will be used, and evaluation of these factors is part of the ISAC assessment. We also consider whether the applicants have experience of using the relevant linkages or, if not, have

discussed them with CPRD and provided evidence of that in the application. The use of linkages is also a key part of the confidentiality risk rating system (see below).

Confidentiality risk assessment

Since mid-2013 all studies have undergone a confidentiality risk assessment according to a rating system devised in conjunction with the Confidentiality Advisory Group (CAG) of Health Research Authority which categorises proposals as low, medium or high risk. The CAG risk rating is performed independently by both the chair and scientific secretary using a scheme which takes into account:

- (a) the number of datasets to be linked (1-4 = low risk; 5-7 = medium risk 8+ = high risk) and whether they are established or bespoke
- (b) whether data fields will be used which have a potential or identified risk for re-identification
- (c) whether the investigators have access to any of the linked datasets in a patient identifiable form, or associated with a patient index
- (d) whether small or specific populations will be studied with a risk of small cell counts based on rarity of outcomes or exposures, and proposed stratifications.

Broadly, the risk category is initially defined on the basis of (a) and then moved up a level if one of the factors (b)-(d) is present. If more than one such factor is present, then it is rated high risk. When the final CAG risk rating is low, no action is required. When it is medium risk some mitigation is required (e.g. reminding investigators of the risk and their responsibilities) and is included in the feedback. When it is high risk,

ISAC completes its assessment of all other issues (if necessary requesting a resubmission) and then refers the protocol to CAG for their approval.

Much the commonest reason for application of the medium risk CAG category is the likelihood that small cell counts will arise and need to be suppressed. A standard statement is used in the feedback indicating that "it is essential that consideration is given to preserving confidentiality at the reporting stage. The possibility of unintentional (deductive) disclosure arises when cells with small numbers of patients are quoted. Please note that, when reporting the data, CPRD policy is that no cell should contain <5 events."

The nature of the study and the potential benefit of the study to public health/interest

The nature of the study is considered by asking applicants to tick one or more of 8 specific boxes on the application form, as follows:

- Adverse Drug Reaction/Drug Safety
- Drug Utilisation
- Disease Epidemiology
- Drug Effectiveness
- Pharmacoeconomics
- Methodological
- Health/Public Health Services Research
- Post-authorisation Safety

If the study is accepted as falling into one of these categories, it is automatically considered to be of potential benefit to public health and in the public interest. However, if the study type does not fit into one of the accepted categories, there is an "other" box and, in such cases, further explanation by way of a statement of the potential benefit of the study to the public health/interest is required and is placed within the lay summary.

Acceptability of the plans for disseminating and communicating the study findings

ISAC expects most studies that it approves to be published in the scientific literature and considers it an ethical obligation for any study with potential public health implications. For applications which do not indicate that publication is intended, we may encourage publication in the feedback or even request a commitment to publish be added to the protocol.

Current protocol review system

In mid-2015, in response to increased volumes and the non-viability of the chair role on a half time basis, a new review system was introduced whereby the assessment of low-risk protocols is performed by members of the CPRD research team and feedback drafted by the scientific secretary with chair oversight and sign-off. The chair continues to handle high-risk protocols as before, preparing the feedback on the basis of reviews from two members and an assessment performed by the scientific secretary. The new review process is summarised in Figure 1.

Outcomes of the application process

The product of the ISAC review process is a feedback document which sent to the applicants with one of four possible outcomes:

(1) Approve

This is recommended for applications which raise no significant scientific or other concerns and appear to have met all ISAC requirements.

(2) Approve with comments

This is recommended for applications which meet the approval criteria given above when there are specific suggestions to improve the research that can be made or minor errors that the applicant ought to correct. In general, this option is only used when such changes that might be anticipated from the comments would not lead to an amendment being considered necessary (the criterion for which is that it would "substantially change the study design or analysis plan" - see below).

(3) Resubmission

This is recommended when it is considered that there are specific significant issues which need to be addressed before the protocol can be considered scientifically acceptable and/or to have met ISAC's requirements. A list of points for the applicant

to address accompanies this recommendation. Resubmissions are assessed promptly and the application approved providing all the points made are considered to have been addressed and the protocol amended accordingly. A small proportion of applications require second and, very occasionally, third resubmissions.

(4) Rejection

The possible scientific grounds for ISAC rejecting an application are as follows:

(a) the study is considered unfeasible in CPRD.

(b) the application is considered to be grossly deficient from a scientific perspective such that, should the applicants wish to pursue the research, it would be preferable to start with a new submission than to try to deal with the concerns by resubmission.

(c) there are major concerns that the research is likely to lead to a flawed outcome.

In 2014, resubmission (59%) and approval with comments (28%) were much the most common outcomes with 11% approved without comment and only 2% of applications being rejected.

Amendments to protocols

In mid-2012 ISAC introduced guidance to applicants relating to the submission of amendments which recognises that, during the course of some studies, it may become

necessary to deviate from a protocol which has been approved by ISAC. In our view, any deviation should be clearly documented by the applicant but not all such amendments need to be submitted for ISAC review and approval. The general principles to be applied in regard to the need for submission are as follows:

- Major amendments should be submitted
- Minor amendments need not be submitted (but must still be documented by the applicant and should normally be mentioned at the publication stage)

We consider an amendment as major if it "substantially changes the study design or analysis plan of the proposed research". The guidance we have developed contains examples of amendments which might be considered major or minor and the current version can be found on the CPRD website [8].

Transparency

Whilst our processes have been open and subject to standard UK government practices (e.g. publishing an annual report) since the Committee was formed, prior to mid-2015 all information relating to specific CPRD protocols submitted to ISAC was held entirely in confidence. Since then, a more transparent approach has been adopted whereby the lay and technical summaries are being published on the CPRD website three months after the study commences [9]. Aside from the general desire to be more open, one of the motivating factors here has been increasing duplication of studies within the database. Some duplication is hard to prevent and may be desirable. However, the knowledge that one or more studies of a particular issue are already ongoing should help to deter the research community from developing further similar proposals in favour of other research questions as yet unstudied in the database.

Discussion

The scope of UK primary healthcare data for addressing important medical research questions is potentially enormous and covers many areas of clinical medicine and public health, resource utilisation and health economics, and methodological studies. The formation of CPRD through linkage of primary and secondary care records, mortality and other databases/registries in recent years has extended the potential value of the data considerably, as demonstrated by an extensive bibliography [10]. However, it is essential that there is an adequate oversight mechanism in place [11] and this is the function of ISAC which provides independent review of research proposals, and whose approval is needed before the research can begin.

The risk-based process we have described has demonstrated the feasibility and acceptability of a robust and transparent process of oversight. Over a long period of time, there have been no important breakdowns of governance or significant public concern about the research being undertaken in CPRD. This is despite the sensitivity of the data and the development of extensive linkages of various databases.

While this is a report on a single system, we believe similar approaches to database oversight exist elsewhere in the UK, and more widely, although there is little published information about their processes. For example the Secure Anonymised Information Linkage (SAIL) databank in Wales has an Information Governance Review Panel (IGRP) which reviews all proposals to use SAIL data "to ensure that they are appropriate and in the public interest", a process which normally takes about three months [12, 13]. Research proposing use of The Health Improvement Network (THIN) database is also subject to review by a Scientific Review Committee [14].

After 10 years, ISAC is facing a rapidly changing environment and is likely to need to adapt the way it operates in the future. The future challenges are likely to include further linkages, new sources of data (e.g. provided by patients) creating different governance challenges, and possibly changes in data regulations relating to health research. There is also growing uncertainty about exactly what is “health” data which could raise different concerns about use and access. We are confident that the principles of oversight discussed above will be maintained and built upon in order to ensure that the quality of research using the database is maintained, and that it will continue to be used in the public interest, and in the best interests of public health.

Acknowledgements

We thank all members of the Committee (listed below) and its Secretariat who have contributed to the work of ISAC over the past 10 years. Particular thanks are due to the scientific secretaries to the Committee: Tarita Murray-Thomas (2006-2012), Kendal Chidwick (2012-4) and James Ellis (from 2015). Members of ISAC: Jennifer Adgey (Chair 2008-10), Krishnan Bhaskaran, Benjamin Cairns, Corinne de Vries, Richard Donnelly, Christopher Edwards, Stephen Evans, Brian Gennery (Chair 2006-7), Martin Gulliford (Acting Co-Chair 2010-11), Ian Harvey, Peter Helms, Amrit Hungin, Iskandar Idris, David Irvine (Deputy Chair 2014-5), Umesh Kadam, Wendy Knibb, Benjamin Lipsky, Paul Little, David Lovell, Sally Malin, Richard Martin, Emily McFadden, Barbara Meredith, Sarah Meredith (Deputy Chair 2006-11), Simon Mitchell, Keith Neal, Barbara Pierscioneck and Ruben Thanacoody.

Conflicts of interest

None relevant.

References

1. Lawson DH, Sherman V. & Hollowell J. for the Scientific and Ethical Advisory Group. The General Practice Research Database. *Q J Med* 1998; 91:445–452.
2. Herrett E, Gallagher AM, Bhaskaran K et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int. J. Epidemiol* 2015; 1–10.
3. Yellow Card Scheme. <https://yellowcard.mhra.gov.uk/the-yellow-card-scheme> (accessed 18/11/16).
4. ISAC section of the CPRD website. <https://www.cprd.com/ISAC> (accessed 18/11/16).
5. Minutes of ISAC meetings 2012-2015. <https://www.cprd.com/ISAC/Minutes.asp> (accessed 18/11/16).
6. Annual reports of ISAC.
<http://webarchive.nationalarchives.gov.uk/20141008011653/http://www.mhra.gov.uk/Committees/IndependentScientificAdvisoryCommitteeforMHRAdatabaseresearch/Annualreports/index.htm> (accessed 18/11/16).
7. ISAC Protocol guidance for applicants.
<https://www.cprd.com/ISAC/protocolguidance.asp> (accessed 18/11/16).

8. Guidance on amendments to protocols approved by ISAC.

<https://www.cprd.com/ISAC/amendments.asp> (accessed 18/11/16).

9. Approved studies using CPRD data. <https://www.cprd.com/ISAC/datause.asp>
(accessed 18/11/16).

10. Bibliography of research papers using CPRD.

<http://www.cprd.com/Bibliography/Researchpapers.asp> (accessed 18/11/16).

11. Tene O & Polonetsky J. Beyond IRBs: Ethical Guidelines for Data Research.

<https://bigdata.fpf.org/papers/beyond-irbs-ethical-guidelines-for-data-research>
(accessed 18/11/16).

12. SAIL Databank Governance. <http://www.saildatabank.com/governance> (accessed
18/11/16).

13. SAIL Databank Application Process. [http://www.saildatabank.com/application-](http://www.saildatabank.com/application-process)
process (accessed 18/11/16).

14. IMS Health Ethics. <http://www.epic-uk.org/our-data/ethics.shtml> (accessed
18/11/16).

Legend to Figure 1 (see separate jpeg file).

ISAC review process for new protocols introduced in July 2015.

Table 1: Scientific Assessment Form

Short title of protocol:	Date:
Protocol number:	Name of Assessor:

Section	Protocol	Acceptable?		Comments
		Yes	No	
	Title	<input type="checkbox"/>	<input type="checkbox"/>	
A	Lay Summary	<input type="checkbox"/>	<input type="checkbox"/>	
B	Technical Summary	<input type="checkbox"/>	<input type="checkbox"/>	
C	Objectives, specific aims and rationale	<input type="checkbox"/>	<input type="checkbox"/>	
D	Background	<input type="checkbox"/>	<input type="checkbox"/>	
E	Study Type (descriptive, exploratory, hypothesis testing)	<input type="checkbox"/>	<input type="checkbox"/>	
F	Study Design	<input type="checkbox"/>	<input type="checkbox"/>	
G	Sample Size	<input type="checkbox"/>	<input type="checkbox"/>	
H	Data Linkage (if required)	<input type="checkbox"/>	<input type="checkbox"/>	
I	Study Population	<input type="checkbox"/>	<input type="checkbox"/>	
J	Selection of comparison group(s) or controls	<input type="checkbox"/>	<input type="checkbox"/>	
K	Exposures, Outcomes and Covariates	<input type="checkbox"/>	<input type="checkbox"/>	
L	Data/statistical analysis	<input type="checkbox"/>	<input type="checkbox"/>	
M	Plan for addressing confounding	<input type="checkbox"/>	<input type="checkbox"/>	
N	Plan for addressing missing data	<input type="checkbox"/>	<input type="checkbox"/>	
O	Limitations	<input type="checkbox"/>	<input type="checkbox"/>	

Please provide below any general comments, assessment of issues which relate to multiple sections of the protocol or other information to support your recommendation			
Recommendation (tick one only): <input type="checkbox"/> Approve <input type="checkbox"/> Approve with comments <input type="checkbox"/> Resubmission Required <input type="checkbox"/> Reject			
	Have you a conflict of interest that needs declaring?	Yes <input type="checkbox"/>	No <input type="checkbox"/> Comments: