

# Preserving consent-for-consent with feasibility-assessment and recruitment in clinical studies: FARSITE architecture

John AINSWORTH<sup>a,1</sup> and Iain BUCHAN <sup>a</sup>

<sup>a</sup> *School of Community Based Medicine, University of Manchester, UK.*

**Abstract.** Best practice guidance for clinical studies asks investigators to employ the highest possible standards in privacy and consent. When considering the feasibility of a clinical study, issues of privacy extend not only to actual but also to potential study participants. The consent required to access records to determine whether or not an individual might be eligible to participate in a study is sometimes referred to as consent-for-consent. Some initiatives to enhance the efficiency of study-recruitment could compromise consent-for-consent, for example by inviting a patient to take part in a study without the knowledge of their attending clinician. Through iterative working with experts and examination of protocols we explored a range of scenarios for assessing the feasibility of clinical trials and observational studies, and recruiting participants. The main requirement we identified was to speed up feasibility-assessment and recruitment while preserving the patient-clinician trust relationship that is central to consent-for-consent. We present an appropriate information system architecture, FARSITE (Feasibility Assessment and Recruitment System for Improving Trial Efficiency), and show in principle that faster recruitment into clinical studies need not compromise best practice in privacy or consent. We show that FARSITE is a specific instance of an 'e-Lab' architecture for assembling data, methods and expertise around study protocols and defined populations.

**Keywords.** clinical trial, clinical study, design, feasibility, protocol, consent, recruitment, privacy, ethics, law, e-Lab

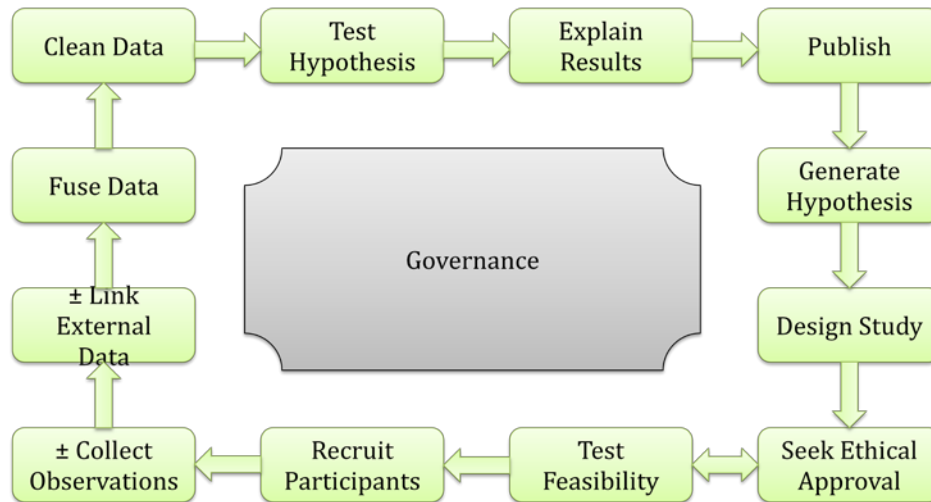
---

<sup>1</sup> John Ainsworth, Health Methodology Research Group, School of Community Based Medicine, University of Manchester, Oxford Road, Manchester, UK. M13 9PL. john.ainsworth@manchester.ac.uk

## 1. Introduction

Medical research has explicit governance in most nations, which has been guided internationally by the Declaration of Helsinki and its revisions over the past 45 years [1]. Some nations and agencies have more time-consuming requirements than others for the administration of medical research, and participation in clinical studies can be difficult to achieve. In a review of UK-supported clinical trials [2] more than half of the investigators asked the funding agency for an extension and a third did not hit their recruitment targets. A recent point of debate in the UK about inefficiency in clinical studies has centred on consent. Specifically, consent-for-consent, which means the consent required to search an individual's health record to determine whether or not they should be invited to participate in a clinical study [3]. Traditionally this has involved clinicians and researchers making judgements based on the study protocol and local circumstances. More recently, however, research ethics committees have started to move away from this opt-out approach to consent to an opt-in system whereby the patient alone, or the patient with the advice of their attending clinician, must first declare their wish to be approached to participate in a study. This move has been accompanied by a fall in participation rates, and concerns have been raised over the biases it might introduce into studies, thereby lowering the quality as well as the efficiency of medical research [4].

A recent review [5] of methods to improve the participation rates of clinical studies did not consider informatics methods, perhaps because only randomised or quazi-randomised controlled trials were included. Yet there are relevant informatics initiatives, for example: 1) UK, US and EU E-Science projects such as PsyGrid ([www.psygrid.org](http://www.psygrid.org)) and Open-CDMS ([www.opencdms.org](http://www.opencdms.org)); 2) "bureaucracy-busting IT" initiatives such as England's National Institute for Health Research Information Systems Programme ([www.nihr.ac.uk/systems](http://www.nihr.ac.uk/systems)); and 3) initiatives building on national healthcare information systems such as the National Health Service (NHS) Connecting for Health Research Capability Programme ([www.connectingforhealth.nhs.uk/systemsandservices/research](http://www.connectingforhealth.nhs.uk/systemsandservices/research)). There is a need for studies of whole-system informatics to support the clinical research cycle (Figure 1):-



**Figure 1.** The clinical research cycle

Here we propose an information system architecture that links feasibility assessment with recruitment for clinical studies. We examine the issues of privacy and consent in such integrated systems. We report on our exploratory prototypes and our future plans for production quality implementation.

## 2. Background

### 2.1. Status Quo: Ad Hoc Feasibility Assessment and Recruitment

Study feasibility is often assessed on an ad hoc basis by asking clinical staff for estimates of the numbers of patients with particular characteristics they might expect to see in a given time period. With short deadlines and difficulties in accurately searching patient records, over-estimation is common [2].

Once a study commences, eligible patients must be identified and approached. This may be via the patient's attending clinician during a clinical encounter or by notification such as a letter. Patients are typically identified by searching databases or paper records, which may or may not be systematic. In addition, clinical teams may deal with a number of concurrent studies, each with specific procedure for seeking informed consent during recruitment. So the process is laborious and ad hoc.

## *2.2. Efficiency need: Automated feasibility assessment and recruitment*

In order to speed-up studies and reduce selection bias, there is a need for systems of rapid, accurate identification of patients eligible to participate [4,6]. This applies both to feasibility assessment and recruitment. For a given research protocol, the ideal research information system would parse the protocol, form a search query, and enable the study sponsor to assess the potential recruitment in a specific population while varying the inclusion and exclusion criteria. Thus the study design extends into feasibility, which might benefit the design. When the study is approved, the well-understood protocol and search mechanisms are employed in the same e-infrastructure to aid recruitment. Only at the recruitment stage is it necessary to know the identity of an individual patient, and this disclosure can be restricted to the attending clinician and the patient, as required.

## *2.3. Framework for privacy and consent*

The natural framework for issues of privacy and consent is the clinical information governance plus the research governance for any defined population. In the UK, Local Research Ethics Committees tend to map to such populations, as do the commissioners and local providers of care services – in other words ‘local health communities’. Relevant laws, regulations and guidance may operate at a higher level, but the key trust relationships for clinical studies operate locally.

A specific issue for feasibility analysis and recruitment within a general framework of privacy and consent is consent-for-consent. This refers the requirement to seek an individual’s consent to search their *person-identifiable* records to determine whether or not they are eligible to be invited to participate in a study. Draft guidance has been issued for this in the UK [3].

## *2.4. Related work*

Existing reports focus on alerting clinicians and/or researchers to patients who might be eligible to participate in studies. For example Dugas and colleagues [7] report the design and implementation of a workflow system involving email alerts in a tertiary care setting. Weiner and colleagues [8] report an increase in trial participation rates after introducing a paging

alert system for research in an emergency medicine setting. We could find no studies that map to the natural governance setting, which is the population.

Technical advances in distributed system security and semantic technologies for knowledge management have enabled new research in this area [9]. The VOTES project (Virtual Organisations for Clinical Trials and Epidemiological Studies) has prototyped a system based on Grid middleware [10] that can perform distributed data queries across multiple clinical data sets from multiple independent organisations using a dynamic trust model. These queries can be used to find patients matching trial eligibility criteria. The ePCRNs have also developed a Grid based system for distributed data queries [11]. In addition they have developed an ontology-driven query builder system to simplify the construction of complex eligibility criteria.

To date only prototype systems have been reported by VOTES and ePCRNs. However, both systems are technology driven solutions to trial feasibility planning, and do not address the *process* of trial protocol development and trial enrolment employing consent-for-consent. We believe a process-centric design methodology to be crucial for the development of systems used in a clinical care setting; if the system cannot be accommodated into existing processes then it will fail. In a similar way, the NHS adopts a cautious approach to emerging technology and as a consequence systems such as VOTES and ePCRNs that employ Grid technology will face many barriers to adoption.

### *2.5. Our approach: e-Lab integrated*

We envision a system of feasibility assessment and recruitment that integrates fully with the study cycle, merging with study design phases leading to feasibility, and study management phases following recruitment.

We think of information systems to enable research using anonymised personal information from a defined population as electronic laboratories, or e-Labs, bringing together data, data processing methods and expertise in a secure environment. We see feasibility analysis and recruitment as specific use case for a more general e-Lab, within the same framework of

privacy and consent. This fits particularly well with consent-for-consent. For research organisations working across a number of populations, we see a federation of e-Labs, reflecting local trust relationships and consistently interpreting research protocols.

In the following sections we describe the requirements for such a system, its architecture and initial prototype implementations.

### **3. Requirements**

Our requirements capture process started with interviews with clinical experts and protocol development experts from the Local Research Networks and was followed up with iterative design based around user interface and system prototypes.

The primary requirements for the FARSITE (Feasibility Assessment and Recruitment System for Improving Trial Efficiency) system were identified as preserving the consent-for-consent model for clinical trial recruitment; to improve the efficacy of the clinical trial protocol design process; and, with reference to Figure 1, to automate as much as possible of the workflow from “hypothesis generation” through to “recruitment”. We address each of these three fundamental requirements in detail in the following sections.

#### *3.1. Preserving consent-for-consent and patient-clinician relationship*

It is essential that the system preserves: i) the privacy of the patient in respect of queries to identify patients eligible to be invited to participate in studies; and ii) the clinician-patient relationship in respect of protecting patients from inappropriate invitations to participate in studies – for example when a patient is grieving. The system must be flexible to accommodate changing interpretations of consent-for-consent.

#### *3.2. Improving research protocol design interactively*

When designing a research study protocol it is important to balance the need for tightly defined eligibility criteria against the need to get sufficient numbers of participants to achieve the required statistical power. A system that enables the user to progressively test and refine eligibility

criteria is required until the correct balance is found. An automated system that performs a parameter sweep by testing all combinations of all eligibility criteria within bounds and increments specified by the user would be efficient.

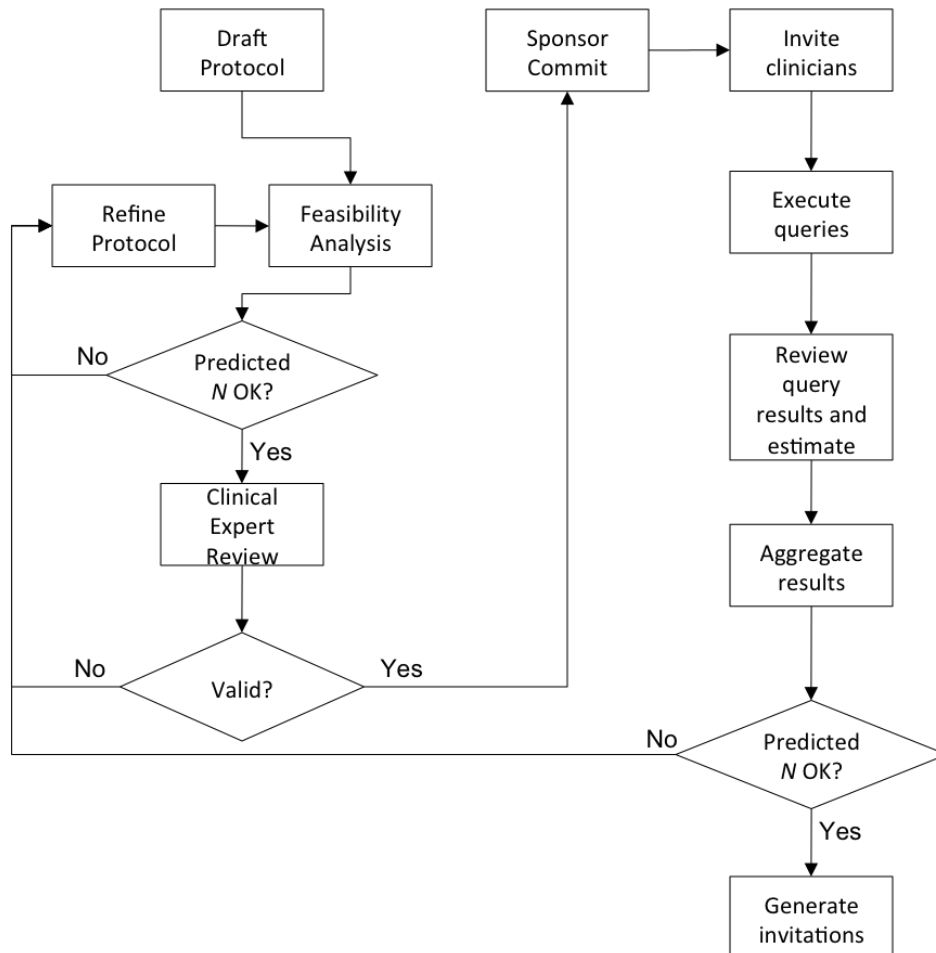
The definition of the eligibility criteria must allow the user to select clinical codes quickly and easily. It must also allow for complex combinations of criteria using Boolean operators. For individual criteria it must be possible to require an exact value or to specify an upper and/or lower bound.

We recognise that some study criteria may not be recorded (accurately) in electronic health records [12].

### *3.3. A unified process model*

Existing approaches [5,13,14] have either focused on the trial protocol design or trial recruitment, without making the connection between the two. We argue that this is essential to ensure that efficiency gains made do not compromise best practice for privacy and consent, and any solution must recognise a clear distinction is required between the actors involved in clinical trial protocol design, and those involved in clinical trial recruitment [11].

We have captured our subject matter experts' views of the ideal trial protocol development and recruitment process as a simple flow chart. This is shown in Figure 2.



**Figure 2.** The clinical study protocol development and recruitment process

The initial step is to draft the study protocol. This defines the bounds for each of the eligibility criteria. The analyst then iteratively refines the eligibility criteria against the information available and the results of queries. The refined criteria are then examined by a clinical expert for plausibility and practicality. Issues such as known miscoding of clinical data may be uncovered at this stage. The iterations converge when analyst and clinician are satisfied that the results are stable and reasonable. The expected number of eligible subjects and the agreed protocol are then submitted to the study sponsor for approval. If the sponsor approves recruitment begins. The attending clinicians of the patients identified by the queries are notified that they are seeing patients who might be eligible to participate in the study. The clinician elects to see which of their



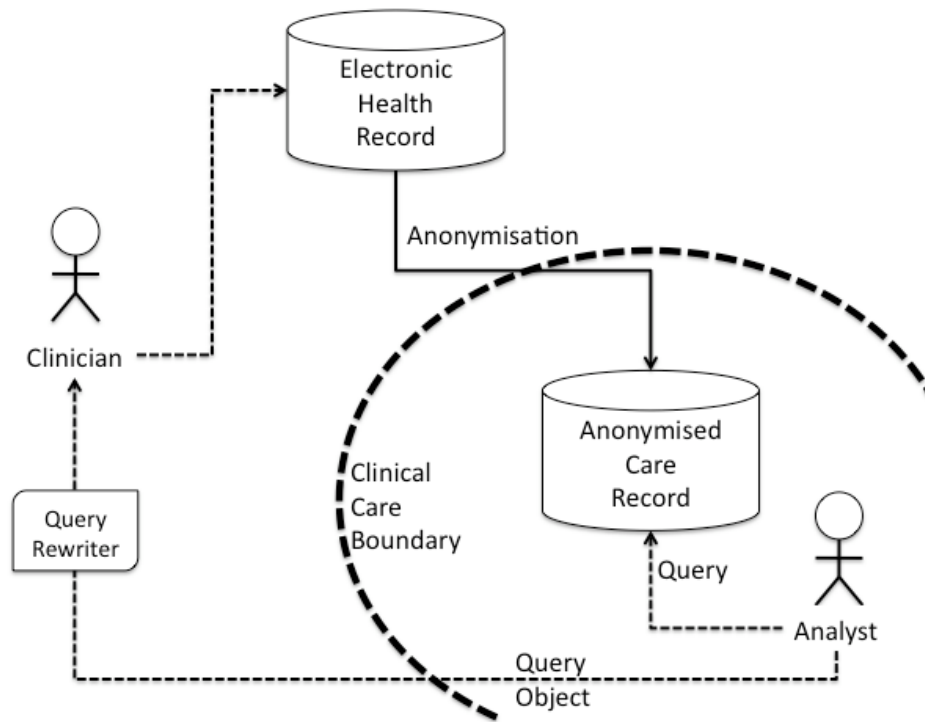
patients meet the eligibility criteria, and makes a judgement about whether or not it is appropriate to invite specific patients to participate. The clinician may notice a coding error that has led to the incorrect identification of eligibility. Or the clinician may feel that a particular circumstance, such as a patient undergoing a divorce, makes it insensitive to invite them to participate in a study. The invitation letters and information leaflets are automatically printed – the clinician may choose to print only for selected patients, or print for all of those patients turned up by the study protocol query and weed out the inappropriate invitations. For some protocols, extra information, not held in the patient record, may be needed from the clinician to complete the assessment of eligibility. As the study progresses, if ongoing recruitment is required, the system must autonomously run the eligibility queries and notify the clinician if any of their patients are found to be newly eligible to invite into the study.

#### **4. Architecture**

Our analysis of the requirements identified the need for the system to distinguish between those who design and coordinate studies, and those who recruit participants. The former are typically researchers or administrators not involved in the direct care of the patient, and the latter are typically the attending clinician of a patient in a context relevant to the study. The clinical researchers should not have unnecessary access to patient-identifiable information – and indeed this is not necessary for their role in protocol design and refinement. They simply need to identify the number of patients that meet a set of eligibility criteria. It is useful to consider the ‘clinical care boundary’ as shown in Figure 3 that divides the clinicians and the researchers. On the clinical side of the boundary there is access to identifiable patient data, whilst on research side there can only be access to anonymised data. Furthermore, no identifiable patient information can cross the boundary. Therefore the system must include an anonymised copy of the electronic health record system that can be queried by clinical researcher as part of the trial protocol design process. The analyst progressively develops the trial protocol by issuing queries against the anonymised repository to determine who many patients will meet the specified eligibility criteria. Although these queries only return an integer count of matching patients, allowing users to issue a sequence

of queries leaves the system open to deductive disclosure [15]. For example, if I know that my next door neighbour has only one leg and is asthmatic, and if the query returns a count of one for "+one-leg +asthma", then I can issue a query of the form "+one-leg +asthma +alcoholic", that will tell me that my neighbour is alcoholic if the count is one. To counteract this, the trial protocol design system will filter the results of queries to ensure that counts less than five are returned as five.

In order to preserve consent-for-consent the attending clinician must run the query that was constructed by the analyst against the electronic health record system to identify potential recruits. This raises a number of technical issues that need to be resolved. Because the anonymised repository does not contain patient identifiable information the identity of the attending clinician is not known. So, the query must be transmitted across the clinical care boundary and then it must be autonomously rewritten to identify all clinicians with potential recruits. Using this information, the query is rewritten again, this time a specific query for each clinician with eligible subjects is created. This query is constructed to return only the eligible patients for whom the clinician is responsible. The query is stored for future execution. The clinician is notified by email that a trial is active and that they have patients that are eligible. The email contains a HTTP link which when clicked on, executes the query for that clinician, assuming that they can successfully authenticate with the FARSITE system. The results of the query are presented to the clinician as a form in their web browser. At this stage the clinician selects the subjects that are suitable and submits this information back to the system. The system collates the responses and the projected number of participants is emailed to the trial protocol designer. This may cause the trial protocol to be redesigned. If the number of participants is acceptable, then another email is sent to the clinicians informing them that trial recruitment can begin. The clinician can then log on to FARSITE through the web browser and generate personalised letters and information sheets for each patient.



**Figure 3.** The clinical care boundary and the information flows in the FARSITE system

Once the trial is registered with the system, the system will autonomously run the query to test if new patients have become available since the last execution.

## 5. Implementation Plans

### 5.1. Trial Protocol Development Tool

We have developed a prototype Trial Protocol Designer tool that enables user to construct eligibility criteria queries and retrieve the number of eligible subjects in an anonymised database of diabetic patients. The web based eligibility criteria builder interface is shown in Figure 4. We have also incorporated the same style of query interface into the openCDMS ([www.opencdms.org](http://www.opencdms.org)) system to enable identification of eligible trial subjects from prospective cohort study data, which enables users to save their eligibility queries.

In response to our prototype our users have suggest improvements to the user interface. They have request that for each one of the eligibility criteria we present the count of patients that satisfy it. From this they can easily see the weighting of each criteria and its impact on the overall total. Planned future developments include the ability to quickly and easily find the correct clinical code, where the system will provide the user with a suggested list of codes based on their eligibility criteria so far and using an ontology to make informed suggestions [16].

The screenshot shows a 'Patient Search' interface. At the top, there is a tab labeled 'Patient Search'. Below it, a header bar contains the text 'How to search: (more...)'. Underneath, there are two radio buttons: 'Match all conditions' (which is selected) and 'Match any condition'. Below these are four rows of search criteria, each with a dropdown menu for the field name, a dropdown for the operator, a text input for the value, and 'Add Remove' links. The criteria are: Hba1c (less than 100), BMI (greater than 12.4), Sex (is Male), and Year of Birth (greater than 1970). At the bottom, there are 'Search' and 'Clear' buttons.

Field	Operator	Value	Action
Hba1c	less than	100	Add Remove
BMI	greater than	12.4	Add Remove
Sex	is	Male	Add Remove
Year of Birth	greater than	1970	Add Remove

**Figure 4.** Prototype of web based Trial Protocol Designer

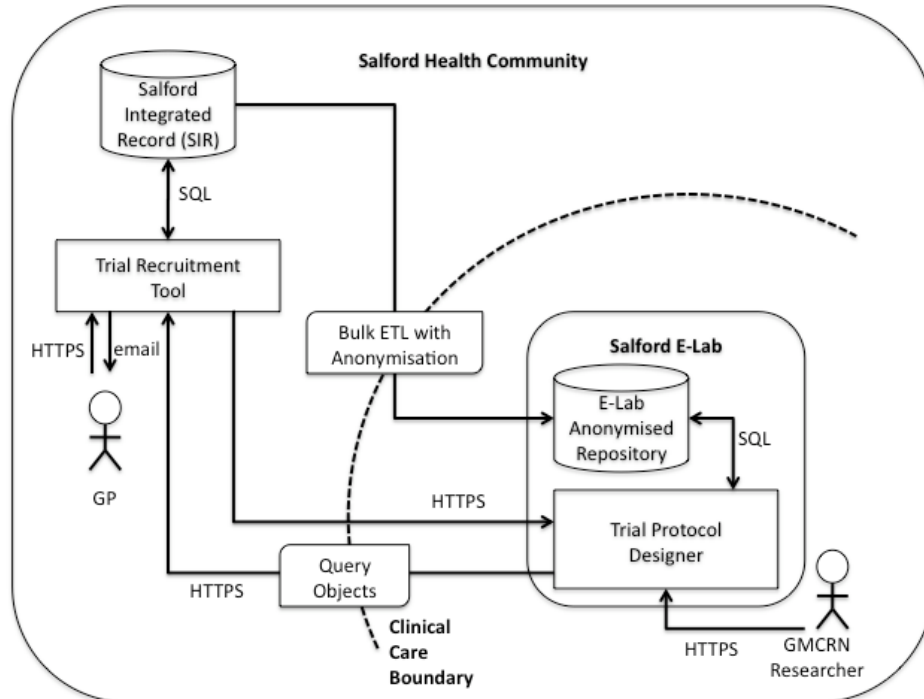
## 5.2. FARSITE in Salford NHS in collaboration with the Greater Manchester Comprehensive Research Network (GMCRCN)

Our first deployment of the FARSITE system will be in collaboration with the Greater Manchester Clinical Research Network (GMCRCN) and the NHS in Salford, and this is shown in Figure 5. Salford NHS has one of the most advance Electronic Health Record systems in England, known as the Salford Integrated Record (SIR), which integrates primary and secondary care data to form a single patient record for each citizen of the city of Salford. Salford is also deploying an e-Lab [17], a secure information system for assembling data, methods and expertise around study protocols and defined populations. The Salford e-Lab contains an anonymised repository of patient data extracted from the SIR system. We plan to mount the Trail Protocol Design (TPD) tool within the Salford e-Lab to enable trial protocol design to be performed by the researchers of the GMCRCN. We will develop the Trial Recruitment Tool (TRT) as a standalone

web application inside the clinical care boundary accessible only to clinicians. The TRT will require two factor authentication, and we hope to be able to use the NHS National Programme for IT (NPfIT) authentication infrastructure which has issues all NHS staff with smart cards. This should enable us to provide a Single Sign On (SSO) solution to the TRT for clinicians. The TRT will interact directly with the SIR system to execute queries on behalf of clinicians to find the identities of eligible patients. The interface between the TPD will be web services using HTTP secured with mutually authenticate SSL.

### *5.3. Integration with HealthGrid*

The model way have presented can be scaled to national or international populations using Grid technologies, whilst preserving the key trust relationship between patient and clinician that operates at a local level. We plan to build a Grid-based federation of e-Labs, initially across the North-West of England, and eventually much wider geographies and populations. OGSA-DQP [18] will be incorporated into the Trial Protocol Designer, and the e-Lab anonymised repository will be exposed through an OGSA-DAI [19] interface, effectively creating a virtualised population-level database, whilst maintaining local ownership and governance. The link between the TPD and TRT will become one to many, but crucially the TRT, and the relationship between clinician and patient that it embodies will be left unchanged.



**Figure 5.** Proposed implementation of FARSITE in Salford NHS for Greater Manchester Comprehensive Research Network

## 6. Conclusion

We have presented a novel architecture, FARSITE, for integrating clinical and research information systems to facilitate the assessment of feasibility and recruitment in research studies, while preserving consent-for-consent. The system is designed to minimise clinicians' work-loads in identifying and recruiting their patients into studies, whilst preserving clinical oversight and ability to protect their patients from inappropriate or insensitive approaches. This and other informatics initiatives to enhance clinical studies should undergo controlled trials as there is a major gap in the evidence base [5]. The need to maximise the population-level utility of health record information, while preserving the privacy of individuals, is not unique to FARSITE – we call the generic architecture 'e-Lab'.

## Acknowledgments

We are most grateful to Dr John New and Dr Martin Gibson of Salford Royal Foundation NHS Trust for their help in leading the clinical activities in identifying the requirements for the FARSITE system.

## References

- [1] R.V. Carlson, K.M Boyd, D.J. Webb. The revision of the Declaration of Helsinki: past, present and future. *British Journal of Clinical Pharmacology* **57**:6 (2004), 695-713.
- [2] A.M. McDonald et al. What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. *Trials* **7**:9 (2006), [www.trialsjournal.com/content/7/1/9](http://www.trialsjournal.com/content/7/1/9).
- [3] Data and Tissues Toolkit: Guidance on Consent for Consent (DRAFT – September 2007). Medical Research Council, 2007.
- [4] J. Hewison, A. Haines. Overcoming barriers to recruitment in health research. *British Medical Journal* **333** (2006), 300-302.
- [5] J. Mapstone, D. Elbourne, I.G. Roberts. Strategies to improve recruitment to research studies (Review). The Cochrane Library, Issue 1, 2009.
- [6] L. Duley et al. Sensible Guidelines Conference: Specific barriers to the conduct of randomized trials. *Clinical Trials* **5** (2008), 40-48.
- [7] M. Dugas, M. Lange, W.E. Berdel, C. Muller-Tidow. Workflow to improve patient recruitment for clinical trials within hospital information systems – a case-study. *Trials* **9**:2 (2008), [www.trialsjournal.com/content/9/1/2](http://www.trialsjournal.com/content/9/1/2).
- [8] D.L. Weiner, A.J. Butte, P.L. Hibberd, G.R. Fleisher. Computerized recruiting for clinical trials in real time. *Annals of Emergency Medicine* **41**:2 (2003), 242-246.
- [9] R. Calinescu, S. Harris, J. Gibbons, J. Davies. Cross-trial query system for cancer clinical trials. In *Advances in Systems, Computing Sciences and Software Engineering: CISSE 2006*. Springer, 2007.
- [10] Stell, R. Sinnott, O. Ajayi. Supporting the Clinical Trial Recruitment Process through the Grid. *Proceedings of the UK e-Science All Hands Meeting* (2006), 61-68.
- [11] K. Peterson, P. Fontaine, S. Speedie. The Electronic Primary Care Research Network (ePCRN): A New Era in Practice-based Research. *J Am Board Fam Med* (2006) **19**, 93-97.
- [12] G. Jenkins, D. Aronsky. A feasibility study for the computerized recruitment of subjects for research studies. *AMIA 2005 Symposium Proceedings* (2005), 996.
- [13] D. Shellie et al. Value of recruitment strategies used in a primary care practice-based trials. *Contemporary Clinical Trials* (2007) **28**:3, 258-267.

- [14] S. Speedie, A. Taweel, I. Sim, T. Arvanitis, B. Delaney, K. Peterson. The Primary Care Research Object Model (PCROM): A Computable Information Model for Practice-based Primary Care Research. *J. Am. Med. Inform. Assoc.* (2008) **15**, 661-670.
- [15] M. Elliot, K. Purdham, D. Smith. Patient Record Data: Statistical Disclosure Control for Grid Based Data Access. *Proceedings of the Second International Conference on e-Social Science* (2006).
- [16] Rector, J. Rogers, P. Zanstra, E. Van der Haring. OpenGALEN: Open Source Medical Terminology and Tools. *AMIA Annual Symposium Proceedings* (2003), 982.
- [17] J. Ainsworth, P. Baker, J. New, M. Gibson, D. Pioli, I. Buchan. Public Health e-Labs: A federated model for e-Epidemiology. *Proceedings of Public Health Informatics 2007* (2007).
- [18] M.N. Alpdemir, et al. "OGSA-DQP: A Service for Distributed Querying on the Grid" in the *Proceedings of the Advances in Database Technology*, 14-18 March 2004, Crete, Greece. Springer, LNCS 2992 858-861
- [19] Chue Hong, N.P., Antonioletti, M., Karasavvas, K.A. and Atkinson, M. Accessing Data in Grids Using OGSA-DAI, in *Knowledge and Data Management in GRIDS*, p3-18, D. Talia, A. Bilas, M.D. Dikaiakos (Eds.), 2007, ISBN: 978-0-387-37830-5.