Reviewer’s report

Title: The SAIL Databank: building a national architecture for e-health research and evaluation

Version: 1 Date: 3 March 2009

Reviewer: John Parkinson

Reviewer’s report:

This is in the main a well written paper that describes the facets of SAIL well.
1. I have no major compulsory revisions

2. Minor Essential revisions
2.1 The abstract does not currently give the population/nation cover of SAIL and this is also not stated clearly or early in the main part of the paper.
2.2 On page 4 the authors’ have not included any statement about the potential for disclosure of any dataset where its population cover is known in a geographic sense when other data, particularly with dated, timed events can be accessed by researchers.

3. Discretionary Revisions
3.1 The dataset has been set up as an "anonymised" dataset and there is no discussion about an ability to validate data that could have been enabled were a pseudonymised approach to have been taken. On the one side this can be taken as a strength, on the other a potential weakness. Perhaps there should be some discussion on this point.
3.2 In the results section there is no mention about data encryption on transfer? This may or may not be relevant.
3.3 There appears no mention of what fields are used in MACRAL.
3.4 In the discussion para 1 the authors might like to review whether it is the work of the OSCHR E-Health Records RB or of the UKCRC-CfH RCP.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I am the Director of the General Practice Research Database and seconded as an expert to the England wide, Research Capability Programme that is also
involved in such record linkage and governance systems.