Reviewer's report

Title: An architecture and method of operation for improving the protection of privacy and confidentiality in population-based disease registers.

Authors: Dr Tim Churches (tchur@doh.health.nsw.gov.au)

Version: 1 Date: 29 Nov 2002

Reviewer: Dr Jules Berman

Level of interest: A paper of considerable general medical or scientific interest

Advice on publication: Accept after discretionary revisions

This is a well-written description of a way of preserving the confidentiality of private data held in disease registers. Though not an implementation, this thoughtful proposal is important, because it forces the reader to envision the problems encountered when handling large amounts of confidential medical information.

This paper should definitely be accepted for publication, but I have a few suggestions that need some response from the author.

1. (minor) In the abstract, the author uses an abbreviation for Disease Register (DR), but abandons the abbreviation throughout the ms. He may as well drop it from the abstract.

2. (minor) He refers to his methodology as something to be used for population-based registers. It seems to me that the proposal is for registers that hold data extracted from patient records. There's nothing about the architecture that requires it to be population-based. The author even describes its use for tissue-registries and biological databases. I'm sure the author has a special interest in population-based studies and may have designed the methodology with this in mind, but there doesn't seem to be any reason to think of the protocol as something that is only suited to population-based registries. So I would suggest that the qualifier "population-based" be removed from the title.

3. (minor) I was well into the paper before I realized that the architecture has never been implemented and that the supporting agencies don't actually exist (identifier translation agency and population register). I think the author should change the wording of the TITLE to indicate that the architecture is a proposal (A proposed architecture,...) and not a done deal. The abstract should also clearly indicate that the proposal has not been implemented.

4. (major). In my mind, the chief problem with the proposal is that each party puts their trust in other parties. On page 13 (section 4), the author writes that the Population Register uses probabilistic record linkage or other methods to find the match. Otherwise it assigns a newly-created id. So you're trusting this agency to do the job right. When errors occur, patients end up with multiple id numbers, and the database becomes corrupted (unique identifiers that are neither unique nor identifiers). This is actually a real problem for hospital information systems and has been well-documented. At least in a hospital, the staff creates their own problems that they themselves need to somehow correct. But when someone else creates problems for you, you're really in trouble. So when you're proposing a
multi-agency architecture, you need to have a way for the participants in the any agency to detect errors introduced by the other agencies and to remedy these problems when they occur. Extending this inter-dependence issue, what would happen, in this system, if there was a successful cyber-attack on an agency that resulted on a re-assignment of id numbers for the different patients, so that a subset of the patients had wrong id numbers? How would the system recover? The author seems to concentrate on discussing confidentiality breaches. But hackers may destroy a system like this without ever breaching anyone's confidentiality.

The author should discuss how to deal with this issue.

5. (strength needs emphasis) In the very last paragraph, the author raises an interesting issue. "To many readers, it might seem unlikely that the level of community consensus necessary for the commissioning of such a system could ever be achieved." Yes, but he answers the objection quite well indicating that the problem requires proposals that can be discussed by the community. That's the real strength of this paper. It serves as a focus for thought and discussion. This should be stressed in the abstract portion of the paper as well.

**Competing interests:**

None declared.