Author's response to reviews

Title: The use of regional platforms for managing electronic medical records for the production of regional public health indicators in France

Authors:

  Marie-Helene Metzger (marie-helene.metzger@chu-lyon.fr)
  Thierry Durand (DURAND@lyon.fnclcc.fr)
  Stephane Lallich (stephane.lallich@univ-lyon2.fr)
  Roger Salamon (Roger.Salamon@isped.u-bordeaux2.fr)
  Philippe Castets (philippe.castets@chu-lyon.fr)

Version: 8 Date: 30 December 2011

Author's response to reviews: see over
Dear Editor,

Thank you for giving us the opportunity to re-submit our manuscript. We tried to answer all suggestions of the Reviewers and refer you to our responses to both Reviewers below.

We hope that the new version of our manuscript will meet your publication criteria and we remain at your disposal for additional information, if necessary.

Sincerely,

Marie-Hélène Metzger

Reviewer's report

Title: The use of regional platforms for managing electronic medical records for the production of regional public health indicators in France

Version: 5 Date: 19 October 2011

Reviewer: Marc Nyssen

Reviewer's report:

- Major Compulsory Revisions

The whole article is messy, the content does not cover the expectations generated by the title.

The article is full of statements that are not supported by data and it is totally unclear which work was actually performed by the authors.

Answer: Ours is a “debate” and not a “research” type of article. The criteria for “debate articles”, described in the journal’s “Instructions to Authors”, are the following: “Debate articles should present an argument that is not essentially based on practical research. They can report on all aspects of the subject, including sociological and ethical aspects”. Manuscript sections for “debate” type articles do not require a “Results” section. However, we added the paragraph 2.6 “Current status of epidemiological platform use for public health in Rhône-Alpes”, to clarify which work was actually performed in the context of the regional platform.

The proposed method of “data mining” is far too vague and has apparently not been tried by the authors themselves.

Answer: We specified which text- and data-mining work is currently in progress (Section 2.6). These activities are undertaken in the context of the ALADIN-DTH project:
- experiments on the extraction and analysis of natural language medical data
- experiments on how to alleviate the frequent characteristics of medical data extracted from EHRs which are primarily designed for patient care
and not for epidemiological research (dataset preparation, diversity of covariates, missing data, class imbalance problem, ...).

So neither on the architecture proposal nor on effective data exploration is this article conclusive, convincing or credible.

**Answer:** The regional architecture proposed is already used for metadata extraction, showing its feasibility. The challenge is to mine free-text medical data in coming years. The feasibility of generalizing text- and data-mining techniques that we are developing in the context of the ALADIN-DTH project will be tested by the regional epidemiological platform in coming months.

- **Minor Essential Revisions**
  Improvement of the English! Many minor mistakes, such as the frequent use of "textual information" should be "text- or text formatted information" ...

**Answer:** The term "textual information" is currently used by scientists in this field, which comprises American scientists (see, for example, reference 1). Moreover, the manuscript was translated by a professional English medical editor/translator (Ovid Da Silva, IRTC, Inc. website: http://www.irtcinc.com/index.htm). Also, the quality of the revision in terms of the English language has again been verified by IRTC, Inc. The term "textual" has been replaced by "free-text".


- **Discretionary Revisions**
  Complete rewrite essential, after the authors have done some real work themselves!

**Answer:** A paragraph, “2.6. Current status of epidemiological platform use for public health in Rhône-Alpes”, has been added to clarify which work has already been done by the authors.

**Level of interest:** An article of insufficient interest to warrant publication in a scientific/medical journal

**Quality of written English:** Not suitable for publication unless extensively edited

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

**Declaration of competing interests:**
I declare that I have no competing interests

Reviewer's report
Title: The use of regional platforms for managing electronic medical records for the production of regional public health indicators in France

Version: 5 Date: 3 October 2011

Reviewer: Daniel J Friedman

Reviewer's report:
This manuscript describes a potentially important French national development in the use of electronic medical records for population and public health purposes. The manuscript is thoughtful, interesting, and well-documented, and should be of substantial interest to international audiences.

Some additional clarifications would improve the utility of the manuscript. These additional clarifications are of four types.

1) First, conceptual clarification is needed regarding whether the regional EMR data will be purely patient-based or truly population-based.

Answer: Actually the platform is based on patient EHRs. However, 2.6 million of 6 million regional inhabitants have a unique regional identifier in the SISRA platform in 2011. The representativeness of this population is discussed in Paragraph 2.6, and Figure 4 has been added. The consequences, in terms of actual development of public health indicators, are discussed. The development of indicators which monitor functional assessment and quality of life is currently preferred. The feasibility of converting it to a population-based platform for prevention purposes will be studied in coming months.

It is not clear from my reading of the manuscript whether the individual national health identifier (INS) will be assigned to each French resident, and whether the EMR will be generated based on the INS or on patient encounters with the health care system, or both.

Answer: The need for a national health identification number has been the object of widespread mobilization of French epidemiologists in the last 10 years, notably when DMPs were created[^10]. Various methods have been proposed[^11]. In its conclusions of February 20, 2007, the Commission Nationale de l'Informatique et des Libertés (CNIL) excluded the national social security number as the national health identification number and advocated the creation of a specific INS to be generated from the national social security number. Article L1111-8-1 of the public health code defines the regulatory framework for the creation of the INS: [translation] “An identification number of health-insurance beneficiaries who are under the care of a
A health professional or medical institution or are part of a health network is used for storing, hosting and transmitting health information. It is also used to open and maintain personal medical and pharmaceutical records." From this regulatory framework, ASIP Santé recently defined the INS development program. An individual INS will be assigned to each health insurance beneficiary for life, and will be non-identifying, that is to say, it will be impossible to deduce any information from this identifier, and knowledge of an INS will not enable anyone to match it to the social security number. This identifier will be randomly generated (INS-A) from a nationally-centralized system for each health insurance beneficiary independently of his/her encounter with the healthcare system. The timetable for implementation of this INS is the responsibility of ASIP Santé and is not yet established to date. In the meantime, to continue rolling out health information systems, a temporary INS has been implemented. This temporary INS, so-called “calculated” INS (INS-C) was introduced in 2010 and is generated by an algorithm taking into account the patient’s social security number, first name and birth date. The INS-C is generated at the first resort to care (outpatient or inpatient). The DMP is then actually generated on patient encounter with the health care system using the INS-C. When the INS-A becomes available, it will be theoretically possible for patients to create their own DMP without encountering the healthcare system but this mode of creation is not defined for now. In Rhône-Alpes, pending the availability of an INS, STIC was created. It assigns a regional unique health identification number. There are no plans for patients to create a regional EHR without encountering the healthcare system in Rhône-Alpes.

Similarly, it is also not clear whether the EMR data are collected only from hospital patients, or from hospital and ambulatory care patients, or from hospital and ambulatory patients and also from additional sites or modes of care.

Answer: The following sentence was added in Section 1.3: The aim of this regional platform is to allow medical practitioners to share electronic health records with other medical practitioners (other public hospitals or private clinics, primary care practices) who need medical information to coordinate patient care.

The differentiation between population-based and patient-based EMRs is essential for understanding whether and how EMRs can be used for population and public health purposes. If it would be helpful, I could provide a couple of citations that may elucidate this point.

Answer: We added the ISO definition of a population health record in Section 2: The aim of creating this epidemiological platform is to build a new information system which corresponds to the ISO definition of a population health record (popHR): “a popHR contains aggregated and usually de-identified data. It may be obtained directly from EHRs or created de novo from other electronic repositories. It is used for public health and other epidemiological purposes, research, health statistics, policy development, and health services management.”

The following paragraph has been added (Section 2.6):
The platform does not yet cover the general population of Rhône-Alpes. Because of this selection bias, the regional epidemiological platform does not yet answer the ISO definition of a popHR for the general population. As detailed by Friedman and Parrish [3], "population-based data on the social determinants of health needed for improving policy-making, program design, clinical care and health professional education" are not yet available... The SISRA platform age structure essentially reflects healthcare utilization, which is the mandatory location for creating EHRs. This limit is important in the selection of indicators to be developed with the tool. The development of a collection system of indicators, which monitor functional assessment and quality of life in patients with chronic diseases (malignant tumours, stroke, asthma), is thus currently most relevant.

However, this limit can be reduced if risk factors or prevention interventions in the general population (followed by the detection of cancer, cancer genetics consultations, monitoring vaccine coverage, the fight against addictions, social determinants, etc.) could be entered in DMPs. Actually, as described above, the objective of DMPs is not clear enough to envisage this kind of use. However, its feasibility will be studied in the Rhône-Alpes region. Extraction of these types of data in regional popHR and linkage with data extracted from EHRs would consequently have to envisage the epidemiological platform as a population-based one, which would permit the study of more epidemiological and public health indicators than is actually possible.

2) Second, terminological clarification is also necessary of at two key terms. The "electronic medical record" should be thoroughly defined, and hopefully compared to the standard ISO definition.

Answer: The term “electronic medical record” was replaced by the term “electronic health record” (EHR) and the corresponding standard ISO definition was given at the beginning of the “Discussion” section: “a repository of information regarding the health status of a subject of care, in computer-processable form”.

The following sentence was added: “The regional platform is a tool that converts non-shareable EHRs to shareable EHRs at level 3 of the ISO definition for EHR (across different EHR locations and/or different EHR systems)”

Also, the "personal medical records (DMP)" should be defined.

Answer: The following paragraph was introduced to define the DMP at the beginning of the “Discussion” section: Following this law, the concept of “personal medical records” (DMP) was introduced with the law of August 13, 2004. It corresponds to the ISO definition of integrated care EHRs. However, some conditions were attached to the concept. The DMP is the property of the patient and not of the health professional, that is, it is the patient who decides to create it and authorizes health professionals to enrich it with documents that they consider relevant to the coordination of care. The patient can mask some
information in his/her DMP if he/she judges that it could be disadvantageous to his/her relationship with the health professional.

The concept mixed 2 different aspects of an EHR: health information exchange between health professionals, labelled as EHRs in the text below, and consumer health records entered by the patient himself/herself, termed personal health records (PHR) in the text below. Confusion between these 2 types of functionalities is a major reason for the difficulty in deploying this tool at the national level in France.

The law of August 13, 2004, in the context of regionalizing health policies, was the next step in promoting the development of RPF-EHRs. The Rhône-Alpes region developed RPF-EHRs but without mixing the 2 aspects of the DMP. The system is based on the principle of an EHR for sharing medical information between health professionals. According to the law of March 4, 2002, patients can consult their EHR but cannot manage it themselves.

Will the regional EMR data be linked for each patient over-time? I believe that the answer to this is "yes," but the answer should be clarified in the manuscript. Is the regional EMR data a "virtual EMR," or will the data exist for each patient (or resident?) in a single database?

Answer: The following sentence has been added in Section 2.2: Regional EHRs at the patient level are virtual EHRs in the sense that regional DPPRs store only metadata and facilitate browsing on all local repositories. Conversely, the epidemiological platform (popHR) will generate a single database related to the public health topic studied. Because of the unique patient identifier, it is possible to follow each patient over time and to build cohort studies.

3) Third, clarification is necessary regarding the use of EMRs for public health in several countries. I am more skeptical than the authors of the true applicability of the U.S. references (second paragraph, 3. Feasibility of setting up a regional epidemiological platform from a regional medical records management platform). The U.S. diabetes prevalence and incidence surveillance system was quite geographically limited, as was the example of the 35 Boston hospitals.

Answer: The following sentences have been added at the end of the section 3: However, these experiments were developed in a geographically-limited area or were not representative at the national level in the United States. Generalization of such experiments is subject to a number of barriers that need to be removed. A structured and coordinated strategy of the Division of Integrated Surveillance Systems and Services of the CDC’s National Center for Public Health Informatics, was created in 2005 [Internet site: http://www.cdc.gov/ncphi/] but the “fragmentation of population health data collection and data stewardship responsibilities among federal, state and local governments” stays one of the most important barriers to popHR in the United States.
I also suggest that the discussion of "medical record computerization rates are closely linked to care-system organization and government incentive policies" be slightly expanded (6. Improve EMR computerization rates and quality): of interest is not only EMR computerization rates, but also whether or not those data are aggregated and used for population and public health purposes.

Answer: The following paragraph has been added:

The new French National Convention of General Practitioners and Specialists indirectly endorsed this approach, offering lump sum payments related to public health objectives. One public health objective is the organization of medical practice, particularly in terms of computerization. Article 12.4.1 of the Convention refers to the use of medical records as a tool for coordination of care. Attending physicians must "establish a summary showing the medical treatment plan, including schedule tracking and interaction with other health professionals for advice and coordinated follow-up". It will be interesting in coming years to establish whether these incentive measures are effective for developing meaningful use of EHRs and to evaluate whether these data are practical for popHR.

4) Fourth, clarification is needed of the current status of the use of EMRs for public health in Rhônes-Alpes. I remain unsure of just how many of the 100 health targets can be measured using EMR data: a table, possibly on-line only, listing the targets and indicating which targets can be populated through EMR data would be helpful. The authors indicate that "the use of regional data for public health purposes is under consideration" (2. Design and development of a regional epidemiological platform) should be clarified. For example, will the regional data be used to populate the regional targets?

Answer: We added Paragraph 2.6 “Current status of epidemiological platform use for public health in Rhône-Alpes”, to clarify this point. A table, with pre-selected public health indicators relevant to regional popHR of specific chronic diseases has been added.

Level of interest: An exceptional article

Quality of written English: Acceptable

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

Declaration of competing interests: I declare that I have no competing interests.