Reviewer’s report

Title: The EPIRARE Proposal of a Set of Indicators and Common Data Elements for the European Platform of Rare Disease Registries.

Version: 1

Date: 1 June 2014

Reviewer: jean donadieu

Reviewer’s report:

The authors who coordinated the EPIRARE project, had made a "Proposal of a Set of Indicators and Common Data Elements for the European Platform of Rare Disease Registries."

This is mainly a position paper and not a scientific paper and it is difficult to review it, scientifically, as it is 'proposal'.

Their approach start from a general needs of the EU commission and of the public health actors in the field of rare diseases to have some validated data for rare diseases.

Their wishes are clearly stated: they aim to define a set of data which had to be collected in the perspective of an european data base for rare diseases.

The methodology to propose such list of data is summarized as an expert choices, while the names of the experts are not mentioned and they are far to be co authors of this paper.

The method used by EPIRARE to propose such set of data, may induce confusion. The authors had collected information about registries and have made a synthesis about the information which should be collected by such registries. And then they proposed that not really minimal really a minimal common data set, but the maximum possibilities offers by all the registries.

1) The number of variables of the CDE (common data elements) is higher that those that we collect (I was one of the registry belonging to the epirare survey) and for my best knowledge, it is more ambitious than ANY existing registries. Somewhere it is far to be consensual and it looks more like a list to the Santa’s klaus !

2) the question of the feasibility should at least be mentionned in the discussion...as such data set have never been collected prospectively at my best knowlegde

3) the authors needs to develop ‘for what purpose they want to collect so much information’. The aims of such data collection has to be stated. With such CDE common data elements once can do:

* descriptive epidemiology
* geocodage
* therapy both observatory and therapeutic trial
* correlation genotype phenotype
* quality of life study
* biological research
* situation of healthy carrier (the data base seems to extend enrolment to healthy persons ?)

4) the authors mentioned the orphanet code in order to identify each rare diseases. I am not specialist of all rare diseases, but I guess that I know pretty well one field of rare diseases: the immune and haematological disorders. In this area, based on experience from registry of patients, based on a discussion among registries in this area, the orphanet code appears inadequate in a lot of situations. Several codes represent one single disease, while some disease are not coded. It is also a proposal, not validated and other classifications exist to depict rare diseases: omin, Mesh, Snomed and so far, the synthetis which may represent the ICD11 is not achieved.

SO the question of the classification is not simple... (and the authors in their survey had not asked any question about the orphanet classification)

Again, the proposal is generous, but not validated.
Despite such limitations the authors are very very ambitious:

"The analysis of the indicators requirements showed (see “Additional file 2”) that the completeness of case registration is requested for almost all indicators. Longitudinal observations were required for all those indicators that are related to the patient experience during her/his life."

Beyond the question of the data set, the authors should have to realist and considered that such efforts is very very costly !

As example of such enormous ambitions, the authors had mentioned items like :
*Current and past participation in clinical trials
*Patient willingness to be contacted to participate in a future clinical trial
*Patient willingness to be contacted about donating biological samples
*healthy carrier
*other cases in the family
*healthy carriers in the family
*case parents are consanguineous
*Biomaterial donated
*ID Biobank where the biological sample is stored up
*(if the biobank storing the sample is not known) ID Centre which sampled the biomaterial
*Current off-label drug treatment
*Current drug treatment
*Hospitalizations
*Transplantations
*Surgeries
*Current dietary regimens prescribed as treatment
*Current assistive devices
*Other treatments
*Patient disability profile
*Patient HRQoL index score
*Comorbidity
*Remarkable or unusual symptoms

I strongly propose to the author to evaluate first this proposal by existing registry, before proposing a maximal CDE...

But as a position paper, many things can be said which will only commit the authors!

**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:**

no competing interest