IRDiRC
INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

Meeting report series

Report of the 1st DSC WG on Ontologies and Disease Prioritization teleconference

29 August 2013

Organization

Organized by: Scientific Secretariat
Teleconference

Participants

Dr Peter Robinson, Berlin, Germany (chair)
Prof Michael Bamshad, Seattle, USA
Dr Kym Boycott, Ottawa, Canada
Dr Michael Brudno, Toronto, Canada
Dr Helen Firth, Cambridge, UK
Mary-Beth Harty, Washington DC, USA
Ms Henrietta Hyatt-Knorr, Bethesda, USA
Ms Janine Lewis, Bethesda, USA
Dr James Malone, Cambridge, UK
Dr Ana Rath, Paris, France
Ms Sharon Terry, Washington DC, USA

Apologies

Dr Helen Parkinson, Cambridge, UK

Agenda

1. Presentation of IRDiRC and the mandate of the WG
2. Introduction of the Scientific Secretariat
3. Discussion of issues and concerns for ontologies and disease prioritization

REPORT

Presentation of IRDiRC

IRDiRC is an initiative of the NIH and the European Commission with the purpose of reaching the following two main goals:

- Produce diagnostic tools for most of rare diseases by 2020
- Develop 200 new therapies for rare diseases by 2020

IRDiRC is a consortium of funders investing at least 10 million USD over 5 years in research projects contributing towards IRDiRC objectives and invited patient advocacy group. IRDiRC has three Scientific Committees:

- Diagnostics Scientific Committee (DSC)
- Interdisciplinary Scientific Committee (ISC)
- Therapies Scientific Committee (TSC)

The purpose of the DSC is to try to translate research in rare diseases into diagnostics for patients. The DSC oversees four Working Groups:

- WG on Model systems
- WG on Genome/Phenome
- WG on Ontologies and disease prioritization
- WG on Sequencing

Each WG will bring new insights and strategies to be considered by the DSC. The Chair of the DSC will report the priorities identified by the WG to the Executive. This bottom-up approach will allow better efficiency and will keep the patient at the heart of IRDiRC missions. Recommendations can be provide through a short document (less than 5 pages) that can be circulated to experts for feedback.

This WG has two tasks:

- Define what diseases should be prioritized for research projects and clinical projects
- Define how to develop an unified system for ontologies for IRDiRC to recommend as best use

Presentation of the Scientific Secretariat

The Scientific Secretariat is located in Paris and composed of four full time employees (project manager, communication manager, research scientist and assistant). Employees from Orphanet and the Fondation maladies rares (Rare disease foundation) also provide support in kind. The role of the Scientific Secretariat is to bring organizational support to IRDiRC Executive Committee, Scientific Committees and Working Groups by, among others, helping organizing meetings and teleconferences, writing the report of these meetings/teleconferences, circulate documents and prepare any necessary documents upon request.

Issues and concerns for Ontologies and disease prioritization

Ontologies

Numerous ontologies already exist for the biomedical community and interoperability is essential. Several points about interoperability were discussed:

- What ontologies are needed to boost the research on rare diseases? Disease and phenotype ontologies should be linked. Scientific ontologies? Biobank ontologies?
- What should be the hierarchy of the recommended ontology? Simple hierarchy is easier to set up and less expensive but a more sophisticated interlinked hierarchy would allow more specific and complete query. Orphanet and HPO already have crosslinks. It would also be valuable to have software and tags to navigate ontologies, especially for non-specialists.
- What should be the strategy to reach interoperability? Developing a tool to allow interoperability of existing initiatives, gather experts to build on the best ontology or ???
- Funding for the sustainability of ontologies is a problem.
- How to promote the adoption of the system that IRDiRC will recommend? This is not a topic for this WG. IRDiRC should be able to recommend it as a standard for RD through Global Alliance.

⇨ A review of existing ontologies (disease, phenotype, mutation nomenclature, etc.) and their characteristics (what and how) is needed to identify gaps, avoid duplication of efforts and be able to make a recommendation.

⇨ A global estimation of the cost for ontologies infrastructure and resources sustainability to present to the Executive Committee would be helpful.

⇨ An estimate of the likelihood of solving most RD by 2020 to present to the Executive Committee would be interesting

<u>Disease ontologies</u>

The following points were raised:

- Should RD classifications be interoperable with common disease classifications?
- A clear and precise terminology of diagnostics is necessary as there is a great heterogeneity of diagnosis (clinical, molecular, genetic, etc.)
- Nomenclature of disease is also a problem as it usually evolves from a symptoms/discoverer nomenclature to a genetic nomenclature. The changes in nomenclature should be captured to keep the list consistent with each others.

<u>Phenotype ontologies</u>

Multiple phenotype ontologies with different strengths and weakness already exist: HPO, Orphanet, POSSUM, Snomed, ICD, etc.

Two topics were addressed:

- Minimal set of phenotype terminology: a core phenotype terminology of 2,000 terms will be finalized in Boston in October to improve symmetry of Snomed and ICD systems. This WG

should recommend as soon as possible what position IRDiRC should adopt, i.e., recommend the core phenotype terminology or something beyond considering that 2,000 terms are not sufficient.

- Mechanisms for improvement and update should be developed

Biobank ontologies

The biobank ontologies should be listed but this is not the priority of the WG.

Disease prioritization

Prioritization can have two aims: diagnostics and therapeutics.

Four approaches can be taken to prioritize diseases:

- To get a list of diseases that are currently unsolved.
- To increase tractability to solve disorders that are more difficult to diagnose
- To determine which diseases are not covered by research. This would require an analysis of the research landscape.
- Prioritize diseases that are not linked to a gene and for which genetic testing in clinical testing is inexistent.

The estimation of diseases left to solve is difficult although crossing datasets could help narrowing down a concrete set of diseases.
It is important to increase tractability but solving very rare diseases is also essential. Moreover, there is an ethical imperative to try to identify diseases that could be treated as early as possible.

⇨ The discussion around disease prioritization lead to the conclusion that the WG should focus on group of diseases that are more tractable for a short-term as building dataset and establishing collaboration for shared platform will facilitate solving the diseases that are difficult to diagnose.

Main deliverables

- Provide a list of disease ontologies to the Scientific Secretariat for mapping
- Provide a list of phenotype ontologies to the Scientific Secretariat for mapping
- Provide a list of biobanks ontologies to the Scientific Secretariat
- Provide a list of databases on research projects to the Scientific Secretariat
- Suggest terminology for the different types of diagnostics
- Circulate a doodle to fix the date of the next teleconference to be held in 3-4 weeks