

Meeting report series

Report of the 1st ISC Working Group on Data Sharing and Bioinformatics teleconference

July 15, 2013

Organization

Organized by: IRDiRC Scientific Secretariat
Teleconference

Participants

Prof Matthew Bellgard, Perth, Australia
Prof Tony Brookes, Leicester, UK
Prof Angel Carracedo (WG Chair), Santiago de Compostela, Spain
Dr Matt Hurles, Cambridge, UK
Prof Jeff Krischer (WG co-Chair), Tampa, USA
Dr Barend Mons, Leiden, The Netherlands
Prof Debbie Nickerson, Washington, USA
Dr George Patrinos, Patras, Greece
Dr David Salgado, Marseille, France (invited by Hanns Lochmüller)

Ms Roseline Favresse, IRDiRC Scientific Secretariat
Ms Sophie Höhn, IRDiRC Scientific Secretariat

Apologies

Dr Ivo Gut, Barcelona, Spain
Prof Jacek Majewski, Montreal, Canada
Mr Jeremy Manuel, UK
Prof Eric Boerwinkle, Houston, USA
Prof Hanns Lochmüller, Newcastle, UK
Dr Rachel Richesson, Durham, USA
Dr Andreas Zankl, Queensland, Australia

REPORT

Introduction of the Scientific Secretariat

Following the welcome of the participants by the chairs of the WG, the Scientific Secretariat was briefly introduced. Through this Scientific Secretariat, Orphanet and the French Foundation for rare diseases are bringing 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, and preparing any necessary documents upon request.

Issues and concerns for discussion

Scope and focus of the WG

- ▶ Some confusion amongst WG members about the focus and scope of the WG: “datasharing” and “bioinformatics”. Is it about standards for datasharing? Is it a bioinformatics group looking at datasharing?
- ▶ “Sharability” is recognized as a better-suited word than “sharing”, the issue is not about “sharing the data” but “making the data sharable”. Bioinformatics is about data sharability. The WG is more about data utility than standards about sharability.
- ▶ Priorisation is needed: the WG can’t be in charge of all issues and it has to be pragmatic, starting with use cases.
- ▶ Ethical issues related to data sharability. Question of the minimal number of variants by which you can identify a person. A more global reflection is connected to the WG. Link with the Ethics and Governance WG on this topic?
- ▶ The focus of the WG should be further refined and clarified in order to create something that is achievable. Technical sharability of data once they exist is a key issue.
- ▶ The process to feedback information towards the ISC should be clarified.
- ▶ Clarification is needed with regard to the WG objectives, its specific tasks and the fact that there is no dedicated funding.
- ▶ For some WG participants, the relations with the Human Variome Project (HVP) as well as the Global Alliance should be questioned. Nevertheless, it is stressed by others that at the HVP level, there are no standards determined and developed so far and the Global Alliance only starts.
- ▶ A list of major areas of focus for this WG should be prepared and developed.

Standards being developed or implemented: inventory and awareness of all efforts on-going in the field are needed

- ▶ Putting efforts of this WG into perspective is stressed as a critical step. Major and numerous efforts are undertaken in the US and at national governments’ level as well as worldwide. Standards do exist but are not necessary used.
- ▶ There is a primary need for inventory and awareness of the efforts. Sharing the knowledge of other efforts between WG members will enable to develop such inventory.

- ▶ It was suggested starting by reviewing the standards. This is a large field but efforts can be concentrated on large-scale projects such as RD-CONNECT, Gene2Phene, etc.
- ▶ One first effort would be to look at major projects and national standards with regard to data sharing.

Identifying gaps for data standards

- ▶ Ontologies used depend on the research focus and what is encompassed.
- ▶ Some areas require a stronger interest.
- ▶ Completeness issue: the WG shouldn't try to cover everything. Some projects might be aware where there are gaps and may contribute to fill them.
- ▶ Rare diseases have additional requirements that should be analysed.
- ▶ Phenotype is also stressed as a big issue.

“Recommending” standards

- ▶ Some concerns have been expressed with regard to the risk of a top-down approach.
- ▶ WG members agreed to make recommendations that will not be mandatory. The principle of endorsement would be a relevant level of involvement.
- ▶ Question of who the WG is trying to help? The WG' aim is to help IRDiRC 2020 goals. Data standards are one of the requirements to achieve those goals. Defining standards will be helping everyone working under the IRDiRC umbrella.
- ▶ The objective is to provide a mechanism to encourage a collaborative approach. It is agreed that guiding stakeholders is what remains important.

Access to the data

- ▶ Best practices are mentioned such as publishing the standards to increase the level of their implementation.
- ▶ There is not only an issue of data sharing but of optimizing data access. That's what the Global Alliance comes about to tackle this issue the proper way.
- ▶ Whereas the Global Alliance is mentioned to be a very Anglo-American initiative for the moment, its objective is to get other countries involved soon. The Global Alliance won't change the standards themselves. It goes far beyond the bioinformatics issue itself (The Global Alliance White Paper will be sent to WG members).

Measurable outcomes

- ▶ Reference is made to a RD-CONNECT jamboree meeting that took place earlier in July in Barcelona. On this occasion, discussions were held on IRDiRC and its overarching, ambitious goal. The issue of measurable outcomes was stressed. Examples of measurable patient outcomes discussed included: time to diagnosis, time to treatment, evaluation of treatment, accuracy of diagnosis, etc.
- ▶ A step forward for this WG could be to define what is measurable by focusing on 1 to 3 items and looking at instruments that are well documented and clinically relevant (taking into account that it has been very disease specific so far).
- ▶ This more pragmatic approach (rather than looking at the overall goal) could lead to more tangible outcomes.

Connection with the EU-funded RD-CONNECT Project (FP7)

- ▶ Once the objectives of the WG are defined, contacts will be developed with RD-CONNECT since it appears as one major initiative with regard to data sharability. This will be the opportunity to see to which extent RD-CONNECT want to liaise.
- ▶ It is stressed that strengthening the transatlantic connection may be needed and welcomed since RD-CONNECT is a EU Project.
- ▶ At the RD-CONNECT jamboree earlier in July, SOPs have been defined for omics data. The feedback will be shared with the WG members.

Other connections and initiatives mentioned

- ▶ For the phenotypic ontologies, an IRDiRC meeting is mentioned at fall. A brief description of the IRDiRC WG looking at ontologies might be useful.
- ▶ With “PhenoDB”, what could be the interactions?
- ▶ With regard to “Clinicaltrials.gov”, efforts have been undertaken that require data sets associated to publications (as online appendixes). For the moment, it doesn’t include phenotype data, only genomic or outcome measures.
- ▶ As far as “DBgap” is concerned, there is no connection with data standards and data sharability. Nevertheless, should connections be developed?

Other issues

- ▶ Financial resources: it was stressed that it would be difficult to do something substantial without funding.
- ▶ Different questions were raised on human resources: are there some available? What would be the support from other groups? Data sharability is not something that can be done at an individual/single group level. What are the available resources to achieve the WG objectives within the frame of “IRDiRC projects”?

Main deliverables

- ▶ Collect short bios from all WG members
- ▶ Contact members unable to attend the teleconference to inquire about other issues they want the WG to discuss in September
- ▶ Vision: a one-page document to be developed and attached to the minutes
- ▶ Come back with suggestions on the overarching goals of the WG
- ▶ Follow up confcall with RD-CONNECT
- ▶ Send a doodle to plan the next teleconference to be held
- ▶ Enquiry if there are other opportunities to meet again
- ▶ Circulate the list of WG Participants email addresses