THE HUMAN VARIOME PROJECT

sharing data · reducing disease

an NGO Official Partner of UNESCO

MINUTES

Gene Disease Specific Database Advisory Council
29th July, 2014
9.00 PM Melbourne, Australia time, 1.00 PM Leiden, Netherlands time
by Calliflower

ATTENDEES: COUNCIL MEMBERS

Arleen Auerbach
Nenad Blau
Raymond Dalgleish
Johan T. den Dunnen
Finlay Macrae
Peter Taschner (Chair)
Yves Sabbagh
Bing Yu

ATTENDEES: NON-COUNCIL MEMBERS

Rania Horaitis
Heather Howard
Helen Robinson

AGENDA:

1. **Welcome**

Meeting opened – 9.03 PM (Melbourne time)

2. **Apologies**

Paola Carrera, Richard Cotton, Rosemary Ekong, Pascal Escher, Bruce Gottlieb, Daniel Hampshire, Sarah Leigh, Sue Povey, Mauno Vihinen

3. Welcome new Members

No new members to welcome.

4. Confirmation of minutes of previous meeting – 23rd May, 2014

All agreed, minutes accepted.

5. Matters arising from the previous meeting

5.1 Activity proposal for ethics checklist for LSDBs to be made

Rosemary Ekong will lead this initiative. She has invited others to participate in this group. Mauno Vihinen is confirmed as participating in the group.

Peter Taschner stated that he would write to her to find out more about the progress and if this group has been properly formed.

Finlay Macrae stated that he has a potential Masters student developing an LSDB and who will be looking at an ethical overview of databasing within at least the Australian framework, so he could put some challenges up. Peter Taschner will let Rosemary Ekong know about this.

5.2 Merging existing lists of LSDBs

There are three used lists; one of which is the Gen2Phen list that is not actively maintained. The others are the HGVS list and the Leiden gene variant database list (http://www.lovd.nl/LSDBs) which has all the LSDBs not just those on LOVD platforms and can be queried from your browser using *GENESYMBOL.lovd.nl (e.g., DMD.lovd.nl)*. Leiden is maintaining two lists: one with LOVD databases and one with all LSDBs. ClinVar are also interested in having one up to date list and wish to collaborate in this effort. Part of a proposed merging and collaboration effort will be gathering all the data collected in the past by the different lists and new data envisaged to be useful, so the number of fields may need to be extended to accommodate this in a new definitive list. This will be an ongoing activity. Therefore there should be no need to duplicate efforts and have them make another list. Therefore merging of existing lists should be done immediately.

5.3 Absorbing legacy data into new LSDBs

New LSDBs being formed will be absorbing legacy data from old inactive databases.

5.4 Increase awareness of journals to nomenclature and data sharing

Peter Taschner found out that many of the journal editors are listed as following the recommendations from the International Committee of Medical Journal Editors (ICMJE, http:// http://www.icmje.org). The list can be found at http://www.icmje.org/journals-following-the-icmje-recommendations/. ICMJE is not checking compliance and it is not clear how many journals are enforcing the recommendations. He has contacted the ICMJE asking them to make sure that not only

nomenclature but also data submission statements be included in the journal instructions to authors. The suggestion will be discussed in their meeting in autumn this year.

Raymond Dalgleish mentioned a paper that was published in *Clinical Genetics* where half of the variant descriptions were documented incorrectly in some way. He contacted the Editor, Michael Hayden, to complain about the journal's lack of standards and that they were not accurately reporting disease causing sequence variants. He was pointed to the statement they have in their instructions to authors that seemed to concentrate on CNVs etc.; when it comes to all other data the only statement that they have is "authors are also urged to submit additional information to suitable databases" with no guidance whatsoever. Michael Hayden asked Raymond Dalgleish to provide suitable guidelines for the journal. He had previously directed him to the author instructions for *Human Mutation* that are quite complete. Raymond Dalgleish contacted Mark Paalman, the Managing Editor of *Human Mutation*, to ask if it was ok to plagiarise these guidelines and agreement was given. Are there any guidelines that could be sent now to the particular journal directly?

Peter Taschner stated that he had prepared some letters that need checking. He will send these to Raymond Dalgleish to look over and send something to Michael Hayden.

Subsequent to meeting, Raymond Dalgleish compiled comprehensive variant submission guidelines based partly on the *Human Mutation* guidelines and an draft version has been sent to Michael Hayden for comment. Minor revision have been made by Peter Taschner and the guidelines are ready to be reviewed by other members of the Council.

Johan den Dunnen stated that he had been in contact with Elsevier through their journal *Artherosclerosis*. They have very specific author instructions that mention HGVS variation nomenclature but in addition they also have a nice list of databases that they think are good databases to submit data to. He is looking through the list available at http://www.elsevier.com/about/content-innovation/database-linking#supported-data-repositories. They also have an option for papers in the journal to have an immediate link to the databases. This works through the use of a DOI and he explained that it is very simple using the DOI to have a direct link to LOVD databases so that when someone clicks on the DOI link they immediately have all the variants from that paper in the database; this is very attractive to the publisher so he is looking at the details of this now. This is of benefit to both sides as the paper is also immediately entered in the database. PT: for details and presentations, see http://www.elsevier.com/about/content-innovation/database-linking#about-database-linking

Finlay Macrae suggested that the Council offer to the journals a service that would provide the Editors a facility to provide advice.

Peter Taschner stated that the first step taken is to contact all the journals to find out if they are complying with the recommendations and if not find out why. The next step is to make sure every journal complies with these recommendations, but this is very difficult. The EJHG however is spending some money on personnel to check all the descriptions of variants and to make sure that they are submitted to databases but it is not possible to do this for every journal.

Helen Robinson suggested a “name and shame list”, whereby those journals complying are listed and those who refuse are also listed.

Johan den Dunnen stated that this is not very difficult to do. He also mentioned that he has students that do checks sometimes on journals. They review an entire issue of a journal and then write a letter to the journal stating what is wrong and what they should be doing. The response is usually “how can we police this”. So he suggests that the response back could be that the Council offers these checks to journals for a small fee.

Raymond Dalgleish also suggested that in the guidelines to journals that where a variant has been described in a clinical report that details of that variant should be included as often it is mentioned that the phenotype is caused by a disease causing variant but there is no information at all about the variant.

Action: Activity proposal for ethics checklist for LSDBs to be made

6. Report from Chair

Peter Taschner had a few things to report.

6.1 He received a letter from HVPI mentioning the transition-taking place in the Board. This is now available on the HVP website and was in the HVP transcript earlier in the week.

6.2 He also mentioned that there is a new journal *Human Genome Variation* (www.nature.com/hgv) published by the Nature Publishing Group. It’s more like the Mutation Reports previously published by *Human Mutation*. They are insisting on the submission of data to a database that in this case is their own. He is not aware that they will share the data as the journal has not published anything yet, they are now accepting submissions.

Johan den Dunnen stated that he protested several times that they should not start a new database.

The Editor in Chief is Katsushi Tokunaga; Tony Brookes, Mauno Vihinen are on the Editorial Board and Johan den Dunnen is the nomenclature Editor! The Council should write to them in this regard.

Johan den Dunnen stated that yes, he is listed but if the journal continues with their plan of creating a new database he will not continue with this journal as he does not agree with this.

6.3 Peter Taschner also stated the BRIF Open Journal of Bioresources that publishes database descriptions in very short publications, with the idea that the database is issued with a unique identifier by the journal. They are also publishing descriptions of other types of bioresources. Identifiers might be used to get a research impact factor. It is also a way to tell how many times your database is mentioned in a publication.

7. Gene/Disease Specific Database Activities - Brief report from Council members

Raymond Dalgleish stated that there have been a huge number of new variants, about 600 - 700 added to the *COL3A1* database.

Johan den Dunnen stated that he is now collaborating with a group called PGX for the pharmacogenetics databases; he is discussing with them a proper way for all the genes involved in drug metabolism to come to some kind of standard for reporting the variants and also the haplotypes that they always use.

Arleen Auerbach reported that the Fanconi database is about to add about 100 genomic deletions found by CGH and these are large deletions of genomic DNA.

Finlay Macrae reported that the InSiGHT database has created the notion of "following a variant". It is a chat room faciity and is proving very productive from the point of view of providing enriched data for the database and also for the community. He also reported about the InSiGHT database interface between NGS information with multiple variants of potential relevance to a specific variant and handling that in terms of the clinical activities and handllng in the databae itself.

8. Updates from working groups

8.1 Copyright & Disclaimer Statements on LSDB websites

This statement was open for discussion until a week ago. Peter Taschner looked at it and did not see any reason not to adapt it. The Council needs to wait until the final report that will include all the remarks made by the people responding to this activity. The committee has a number of weeks to incorporate all the comments and suggestions. This should be ready by the next meeting.

8.2 Variant Database Quality Assessment Group

The draft Variant Database Quality Assessment document needs approval by this Council. Peter Taschner has looked at it and is in favour of approving the document in its current text and sending it out for the community to view and comment. He asked other members of the Council for their thoughts.

Finlay Macrae asked if there are any controversial points in the document and the answer was no.

If this document is approved, the next phase is to set up a Database Quality Assessment.

All agreed to accept document.

Peter Taschner will mention this to the Chair, Mauno Vihinen.

9. HVP Global Haemoglobinopathies Initiative

This was brought up by Helen Robinson for information purposes. At HVP5 in Paris in May, it was

agreed to form the above initiative. A small group of core countries – Malaysia, Mexico, Nigeria, South Africa, China, Portugal, Venezuela – have produced a draft project proposal for the initiative. The main aims are to build capacity of HVP Country Nodes in countries where there are large carrier populations of the various genes, and for these nodes to contribute to international disease databases in some harmonised, cordinated and efficient manner over the next 3 to 5 years. The aim is to expand the project beyond the initial group of countries and to raise funds (at national and international levels) to progress the initiative. The project will be overseen by HVP's ICACC, but this Council needs to be aware of the project as well.

Anyone interested in participating or wanting more information, please contact Helen Robinson directly – **hmro@unimelb.edu.au**

10. Recommendations to the International Scientific Advisory Council

The Council recommends that the Variant Database Quality Assessment document be accepted.

11. Other items

Please email Rania (rania@variome.org) if you have any other matters so they can be added to the Agenda for the next meeting.

12. Next Meetings

- September 16th – 9.00 PM Melbourne/1.00 PM Leiden via Calliflower
- November 18th – time TBA, via Calliflower

Meeting Adjourned 10.00 PM