exported to PDF on March 8, 2016

HL7 Clinical Genomics Weekly Call - March 1, 2016

Attendees

1. Bob Milius - NMDP - bmilius@nmdp.org
2. Clem McDonald - NLM - ClemMcDonald@mail.nih.gov
3. David Kreda - david.kreda@gmail.com
4. Larry Babb - lbabb@geneinsight.com
5. Bret Heale - University of Utah/Intermountain Healthcare bheale@gmail.com
6. Joseph Kane - Epic - jkane@epic.com
7. Siew Lam - siew.lam@imail.org
8. Perry Mar - Partners HealthCare System - pmar@partners.org
9. Kevin Power - Cerner - kpower@cerner.com
10. Bob Freimuth - Mayo Clinic - freimuth . robert at mayo . edu
11. Joey Yang - HFUT - jiaoyun@hfut.edu.cn
12. Heming Yao - BCH -
13. Brett Johnson - icanbrj@gmail.com
14. Joel Schneider - NMDP - jschneid@nmdp.org
15. Ning An - HFUT - ning.g.an@acm.org
16. Anwaruddin Mohammad - Oracle - anwaruddin.mohammad@oracle.com
17. Amnon Shabo (Shvo) - Philips - amnon.shvo@gmail.com
18. Jonathan Holt (SeqTech Diagnostics) jholt@seqtechdx.com
19. Siew Lam - Intermountain -
20. Andrea Pitkus - IMO - apitkus@imo-online.org
21. Grant Wood - Intermountain
22. Mollie Ullman-Cullere - Partners -
23. Wanmei Ou

Discussion

- Minutes approval
    - Please take a look at the draft minutes posted here (fyi, I posted a PDF that was exported from the google doc)
    - http://wiki.hl7.org/index.php?title=File:HL7_CG_20160223.pdf
    - Motion to accept - David
    - 2nd - Lam
    - discussion - none
    - abstain- 0
    - nay - 0
    - yea - rest
    - results - motion passes

- 5 min updates
    - ClinGen/ClinVar -
        - nothing new
    - GA4GH -
        - nothing new

- IOM -
  - nothing new
- other -
  - none

- Project updates
  - V2 Lite
    - Proposal: V2 Lite Proposal round III
  - Clem
    - overview
    - 1st division - structural variations, issues are different from rest
    - clarify what we look for vs what we found, e.g., targeted analysis must list targets, or sequence analysis must describe what region is being sequenced
    - small variations discussed today, large variations next week
    - demo of lforms-demo.nlm.nih.gov
    - Bob F. - p notation - determined by exon boundaries, how to choose if there are alternates (multiple annotations per given NM refseq)? Clem - don't know yet, haven't seen an example
    - Bob F - p notation can change over time, one to many relationship
    - Clem - isn't it computable? yes
    - Larry - suggestion, lets do questions at the end. We can get derailed. Clem agrees
    - Clem - if you have an allele id, you can get all you need and autoload
    - Jonathan - reference id? Clem can use oids or identifiers in the future f This will support the "Codeable concept" for future FHIR work and "source" of the code (i.e. Allele ID or transscript ID)
    - Larry: "If" we can standardize on the ClinVar id we should use the "VariantID" not the "AlleleID". ClinVar is moving towards a model where the VariantID can support "simpleAlleles" (as shown here), "haplotypes" (multi-simpleAlleles in cis) and "genotypes" (haplotypes in trans). Every alleleID has a variantID. Structural problem in the ClinVar ID system.
      - Clem: agree
    - Larry: Need to separate the data (gene finding) from the interpretation of the finding (Jonathon agrees). need additional attributes in the "top" part of the report i.e. Disease Assessed rather than the possible disease associated in the "gene finding" section.
    - Clem addressed Chat questions after his presentation. We got to Anwar's question about context, i.e. who is the author of this report (lforms-demo)? who is the audience?
    -
  - Saved Chat
    - 10:01:46 AM from Bob Milius to Everyone:
      - https://docs.google.com/document/d/1jU_DiC933QlPuUIOhdqZOibAcNXCCIn2wouHCZ4LfLs/edit
    - 10:19:55 AM from jonathan to Everyone:
      - How is the reference genome represented? Which one is being used here?
    - 10:22:40 AM from David Kreda to Everyone:

        - But type in questions into chat so they are all collected and can be offered also to Clem afterwards to memorialize responses in his document, etc.
- 10:22:44 AM from David Kreda to Everyone:
    - (Just a thought)
- 10:22:49 AM from Bret Heale to Everyone:
    - Its probably the NCBI's draft at the time the refseek was last updated - But good point Johnathan...also is this the most common splice-variant Or the most 'important' Or the most 'relevant' And what about genes on the other strand...
- 10:23:28 AM from Bret Heale to Everyone:
    - agree with david. if you can, speak and wrtie the question. it will then be avialable for next week (good idea, david k)
- 10:24:30 AM from Larry Babb to Everyone:
    - "If" we can standardize on the ClinVar id we should use the "VariantID" not the "AlleleID". ClinVar is moving towards a model where the VariantID can support "simpleAlleles" (as shown here), "haplotypes" (multi-simpleAlleles in cis) and "genotypes" (haplotypes in trans).
- 10:24:36 AM from Bret Heale to Everyone:
    - Bob F mentioned an issue with the amino acid number being dependent on the refseek choosen and that it can change. Clem answered that it was there as it is common in reports but really is a convienence - change should be computable.
- 10:24:50 AM from Bob Milius to Everyone:
    - This discussion is for small variants - does this mean limited to SNPs? or does this discussion include somewhat larger, like haplotypes of SNPS
- 10:25:22 AM from anwar to Everyone:
    - Question: The amino acid change would depend on the transcipt and in cases if there are multiple transcripts what approach would you use? Show cannonical transtipt or show all of them?
- 10:25:27 AM from Larry Babb to Everyone:
    - We definitely will need to come up with the 2 or 3 minimal variant representations. Currently, ClinVar accepts 2 or 3 and we could start with those since they've been well tested with thousands of submissions from many labs.
- 10:25:41 AM from Bob Freimuth to Everyone:
    - Larry: We need to disambiguate the purpose (whether the intent is to define a location of a variant or a given allele at a given location)
- 10:26:01 AM from Larry Babb to Everyone:
    - we should call "small alleles" "simple alleles" to get on the same semantics as ClinGen, ClinVar?
- 10:27:26 AM from Bret Heale to Everyone:
    - @larry babb. re: 'same semantics'. I agree, small but impt point
- 10:27:44 AM from Bob Freimuth to Everyone:
    - In addition, knowing a variant location and an allele does not indicate observed genotype. We need a way to specify a given # of alleles that as a group indicate patient genotype.
- 10:29:04 AM from Bob Freimuth to Everyone:
    - Labs may not currently list the regions tested, but they should...

- 10:30:25 AM from Larry Babb to Everyone:
    - Clem: in the "Allele" or Gene Findings, if we include the Clinical Significance, we must be able to provide a strict rule of which "disease" or "condition" is the significance context. The possible associated phenotype does not seem like it, if it is it would need a better name. If it is the top level genetic disease assessed then we should be clear that the significance is for all list not just one (in the case that there are more than one).
- 10:31:07 AM from anwar to Everyone:
    - One basic question just to understand the context, who is the author of this report and who would be the targeted audience for this report?
- 10:31:17 AM from Larry Babb to Everyone:
    - I would recommend separating the "finding" from the "interpreation" or "significance" observation. These two observations are typically done be different groups within a lab (one is wet lab, one is geneticist).
- 10:31:34 AM from jonathan to Everyone:
    - Range is tough approach. Some regions of genes have good coverage, but others b/c of pseudogenes.
- 10:31:56 AM from Bob Freimuth to Everyone:
- We need to accommodate preemptive testing (where there is no disease/condition)
- 10:32:01 AM from Larry Babb to Everyone:
    - all "Mutation" terms should be changed to "Variant"
- 10:32:16 AM from Larry Babb to Everyone:
    - that's been an ACMG decision for a while now.
- 10:32:50 AM from Larry Babb to Everyone:
    - We need to be able to pass the genomic dna ref and coordinates (even if labs can't)
- 10:36:28 AM from Bret Heale to Everyone:
    - @Larry Bab. cmmt on: Clinical Signficance. For a testing lab the clinical significance is pre-determined and based on the customer need the test is meeting. So, while I like what you ask for, is this going to be something that fits well into the testing labratory's buisness model. They are not necessarily going to update the specific test panel reporting. But they might make a new one for ordering...just some thoughts
- 10:40:56 AM from jonathan to Everyone:
    - yes, indeed "Allele ID is problematic". Variant ID is better, as long as we represent the reference properly. Agree , remove Allele ID. Cleaner.
- 10:42:11 AM from David Kreda to Everyone:
    - (1) Won't the list of data items will balloon Observation's profile enormously, yet not yet be the clinical interpretations? (This is one reason - not the only one - which makes it quite risky to channel highly variable sized data in Observation. (2) Also if the same data may be repurposed (will be repurposed for sure) to the prior interpretation, how is this a good argument for tight coupling in the same data structure. Post coordinated CDS will need to load and discard possibly the same data multiple times.
- 10:43:33 AM from Mollie to Everyone:

        - Somatic testing is commonly performed in context of ENSEMBL transcripts. Can we add ENSEMBL for all where RefSeq is?
    - 10:43:53 AM from Mollie to Everyone:
        - Agree with Anwar - this is also important for Somatic
    - 10:46:10 AM from Bret Heale to Everyone:
        - by adding interpretation, are you not coupling the data with knoweldge? Isn't data forever and knoweldge evolving? The point being that the interpretation changes but the lab value (bases present or missing) is less variable. right?
    - 10:50:15 AM from David Kreda to Everyone:
        - To Brett's point: the underlying taxonomy of variants can expand hugely and, though it will never approach asymptotically the underlying data (actual sequence), the number will be huge. It is not that ID matching will be the issue - it is ID granting. At some point, why bother, if the search performance will produce rather more interesting results (think Google like search). This is not today's reports, but more or less explicitly discussed in NGS articles.
    - 10:50:21 AM from Bob Freimuth to Everyone:
        - We should clarify the intent of this particular content - is it a structured form of the narrative and intended to be read by humans, or is it purely for computable use? Variant names for humans will need to be aligned with the literature, but that is less important for computers.
    - 10:53:34 AM from jonathan to Everyone:
        - Where is "Allele/Variant Freq"?
    - 10:55:46 AM from Bret Heale to Everyone:
        - many lab tests couple the interpretation with the nucleotide data for single varaint test (in fact, the nuecleotide content is treated as secondary content). this was an effective way to meet customers past needs (they only wanted the interpretation), I believe. do we need to provide a conviencing enough reason to change buisness practices?
    - 10:57:51 AM from jonathan to Everyone:
        - Strongly Agree with @larry !
    - 10:58:12 AM from Bret Heale to Everyone:
        - pardon, I'll work on my spelling.
    - 10:58:35 AM from Amnon Shabo (Shvo) to Everyone:
        - I strongly agree with Larry!!
    - 10:59:24 AM from Mollie to Everyone:
        - does this look like an array of question(disease):answer(pathogenic) pairs
    - 11:00:43 AM from Bob Freimuth to Everyone:
        - Someone went off mute, perhaps computer and phone audio interfering
    - 11:00:48 AM from Andrea Pitkus to Everyone:
        - Curious how the reporting would be compliant with CLIA and where terminology code systems would be added.
    - 11:01:25 AM from David Kreda to Everyone:
        - I would like Clem to address the "beyond report" paradigm in his context setting, not to deprecate the report, but to emphasize the future-purpose CDS
    - 11:02:41 AM from Andrea Pitkus to Everyone:

- right. originatino of the report in a LIS and then downstream uses of the discrete data