HL7 Clinical Genomics Weekly Call - Tues April 5, 2016

Attendees

1. Bob Milius - NMDP - bmilius@nmdp.org
2. Amnon Shabo (Shvo) - Philips - amnon.shvo@gmail.com
3. Jonathan Holt - Seqtechdx - jholt@seqtechdx.com
4. Siew Lam - Intermountain Healthcare - siew.lam@imail.org
5. Bret Heale - Intermountain Healthcare/University of Utah - bheale@gmail.com
6. Gaston Fiore - BCH - gaston.fiore@gmail.com
7. Mollie Ullman-Cullere - mollie.ullmancullere@gmail.com
8. Perry Mar - Partners HealthCare System - pmar@partners.org
9. Bob Wildin - NHGRI/NIH - bob.wildin@nih.gov
10. Gil Alterovitz - gilusa@gmail.com
11. Larry Babb - lbabb@geneinsight.com
12. Terry McDonnell - Syapse - terrym@syapse.com
13. Anwaruddin Mohammad - Oracle - anwaruddin.mohammad@oracle.com
14. Andrea Pitkus-- apitkus@imo-online.com (joined late)
15. Chethan Makonahalli
16. Clem McDonald
17. Elizabeth Newton
18. Eric Whitebay - Epic
19. Heming Yao - BCH
20. Jeffery Karp
21. M'Lynda Owens
22. Robert Barkovich
23. JD Nolen
24. Bob Freimuth

Discussion -

- Minutes approval (Mar 29)
    - http://wiki.hl7.org/index.php?title=File:HL7_CG_20160322.pdf
    - motion/second = Lam/Clem
    - Discussion - Jon clarified his proposal for in email after the meeting, this is included in the minutes; Bob F in vote for minutes, nays should be abstains
    - yea/nay/abstain = 23 / 0 / 0
- Updates from external efforts (GA4GH, ClinGen/ClinVar, IOM, etc)
    - ClinGen -
        - Larry - Nothing new
    - GA4GH
        - Gil - nothing new
    - IOM
        - JD - Nothing new, name changes
    - FHIR

        - Jon's proposed change (refseq id is 1:1) implemented in freeze
        - Clem - is the CG spec in a private service? Gil - All CG products are on hl7 site. The connectathon build is here: http://hl7.org/fhir/2016May/index.html
          We use fhirgenomics site to show various options and how various pieces fit together/educational materials,etc, but then move to current build as approved by CG.
- Draft agenda for Montreal - Due Apr 8
    - Bob Wildin - will be attending for 1st time.
    - Jonathan is looking for funding to attend.
- 0 vs. 1 based nucleotide numbering - continued discussion
    - Clem - favors 1 based; NCBI is storing 0-based, but public face is 1-based
    - Mollie - when submitting to genomic pipeline to repository, 1-based, VCF is 1-based, if we use 0-based then we need clear warning and documentation for implementers.
    - Jon - used to favor 1-based, but now in 0-based camp; if we use this as a low level data structure and separate the data from the observation
    - Bob Wildin - can see both sides, recommends that we err on the side of human readable form. Transport of information is not on computing.
    - Mollie - HL7 shouldn't change what NCBI is doing.
    - Jon - separate the data from the presentation
    - Clem - the message is not the storage
    - Larry - add a type field for 0 or 1 based; objection is too complicated; more difficult to go from 1- to 0- based. Easy to go from 0- to 1- based
    - Mollie - suggested to default to 1-based, but allow for 0- based typing for advanced users
    - Gil - reviewed different use cases, can see everyone's viewpoint, looks like g4gh is 0-based. Also pointed out that BAM (used in many pipelines) is zero base and then the output of the pipeline is VCF which is 1-base. There appears to be current consensus to make v2 1-based. FHIR can deal with it after further information is explored and consider the division of information per Jon Holt idea to divide the underlying data as 0-based and presentation layer as 1-based (e.g. FHIR subgroup).
    - Lam - sees advantages to both, don't restrict usage to one or the other.
    - Larry - we're creating a profile, but others will create other profiles based on this profile
    - Mollie - default should be 1-based, most pipeline tooling is 1-based.
    - Bob M - why need a default? Whats wrong with expliciting saying what type it is? Let's have a required type attribute. Can constrain specific use case profiles to 1 or 0 depending on need.
    - Lam - don't assume tooling won't change
    - Clem - use the most widely used version, don't mix
    - Mollie - can we default V2 on 1-based and FHIR offer both?

    - Mollie - can we get input from NCBI, EBI, COSMIC on this?
    - Bob W - talked with Les Biesecker at NHGRI and they suggest 1-based to start with, as does Marc Williams at Geisinger, at least to start with. INOVA ITMI reflected the bioinformaticians' computational ease argument in favor of 0-based.
    - Lam - 0 vs 1 based not restricted to genomics; mathematicians deal with this too
    - Reach out to other groups
    - Jon - reaffirms his opinion on 0 based; please read the links that were sent out.
        - To clarify my position: I think we need to have a conceptual discussion regarding the purpose of the FHIR Sequence resource. I have been arguing that we think about it as a data-structure that facilitate unambiguous interoperability of the underlying DNA/RNA sequences, not the higher level concepts of observations and/or interpretations. I would argue the value of a "Sequence" resource is as a data structure that the ObservationGeneticsProfile references to facilitate unambiguous computational representations. A 0-based data-structure is able to represent the "fence-post" problem of substrings. My problem is that we haven't had that discussion and at present the current version of the sequencing resource is a proxy for a fancy Observation and not as a data-structure. We need to think about them separately: Data → Observation → Interpretation. V2 and V2 lite belong in observation and at the level of the observation resource and V2, I am in favor of a 1-based representation model including HGVS nomenclature for that purpose as it is more nature, albeit lossy…. Which is why we have a full reference to the underlying data-structure (the sequence resource) that provides the data to "show your work". I can guarantee that if we split and allow both using a "type" attribute, then we will not be able to have true unambiguous interoperability of the underlying data. (Jonathan)
    - Gil clarification on above: Whether or not we choose to go that way, I just want to note that FHIR does allow for this and FHIR subgroup can discuss FHIR-specific issues as well.
    - Bob W - having a field would be good to specify 1 or Zero
- Formation of a DIM subgroup (deferred from last week)
    - Deferred to next week