NHGRI Genomic Medicine III

Sequencing Working Group

Participants: Heidi Rehm, Erwin Bottinger, Howard Jacob, David Ledbetter, Debbie Nickerson, Scott Devine, Kelly Frazer, Scott Weiss, Rex Chisholm, Eric Green, Jeff Schloss, Olivier Harismendy, lucia Hindorff, Zivana Tezak, Josh Denay, John Harley, Wojciech Wiszniewski, Chris Bradburns

End Goal: Return of results from Genome and Exome sequencing for patient care.
Focus on quality not which data are in the reports. The same variant should be called in all clinical laboratories.

Organizing Principles:

- Build metrics for 3,000 clinical genes
  - Clinical context matters with respect to what/how is being looked at—metrics should not
- Build metrics for the remaining genes and rest of genome

Section 1: Wet lab best practices.

Key issue: Laboratories are in need of guidelines for operating platforms.

Solution: CAP/CLIA will develop Expectations for covering the relevant regions based upon the indication for testing (disease gene/locus list, whole exome, whole genome).

Key issue: Quality control metrics and measurable are not consistently defined

Solution: This workgroup will work definitions, metrics, specificity, sensitivity etc. Will work with CAP/CLIA

- Define the metrics that will remove the need to a second method to follow-up.
- Formats need to be defined for variants, etc.

Key issue: Laboratories are in need of standard samples for validating platforms

Solution: This group will develop gold standard samples. Some (all?) will be the same as the ones used by CAP/CLIA

Manufacturers would use these samples

How many samples should there be? Perhaps same seven used for HAPMAP.

Diversity?, complexity?

Action Items:

1. Heidi Rehm to link us to CDC group. Determine which two samples will be used by them for QC. Can these be part of the DNAs to be used by CAP/CLIA.
2. Need to link to CAP/CLIA group
3. Write a white paper about the samples and metrics that can be used to compare sequencing platforms

Section 2: Analytical best practices.

Key issue: Need of a defined set of standards and tools for analyzing genomic data

i. Standards are needed to assess quality (duplicate rates, minimum coverage, quality metrics)
ii. Standards are needed for measuring false positives and false negatives (sensitivity/specificity)
iii. Standards should be platform independent.

Key issue: Need for software, standards, and tools that feed into diagnostic market.

i. Data analysis tools are developing so quickly that it is difficult to define appropriate parameters for analysis.
ii. Software and databases that lock, rather than dynamically change to support the fact that software and processes must be validated.

Solutions:

Data sets that can be used to compare new tools too.
Ways to benchmark software performance. Need to establish these benchmarks

Action Items:

1. Collect a list of data sets, and a bit about them, that could be used for making comparisons. For these data to be part of the program. They must have to correct consent for distribution to other sites.
2. Need to select the Data sets to distribute
3. Define benchmarks

- **Section 3: Central repository for clinical comparisons.**

I. Key issue: Determining the clinical relevance of genetic variation will require large cohorts of well phenotyped individuals, and centralized databases are needed.
   a. ClinVar is one example, but reporting standards are not always clear.
   b. BIC is another example, noting that Myriad stop reporting
   c. Different types of submissions: Observed variants such as in a phenotype to be diagnosed or healthy population.
   d. Large databases are needed to aid interpretation

II. Key Issue: Interpreting actionable variants
   a. What is an actionable variant, how do we deal with it.
   b. Managing Variants of Unknown Consequence
   c. Guidance for lab directors