<u>Variant interpretation breakout group notes from Binfx network meeting 12th June 2017, Leeds</u>
Rebecca Haines (Nottingham), Rodney Nyanhete (GOSH), Phil Davidson (St Georges), Joseph Halstead (Birmingham), Mike Cornell (UoM)

Focus for discussion: Implementation of ACMG variant interpretation guidelines

- There is a need for bioinformaticians to understand the data used, particularly the limitations of data sets (such as the application of ExAC data/constraint scores for late onset disorders)
- Several labs are developing tools for recording the data and calculating the classification using the ACMG criteria
    - Concerns about every lab doing this independently and therefore slightly differently

Action:
Develop a unified data model for storing the results of variant analysis in order to facilitate sharing of the data.
Rodney (lead) and Phil to collaborate on this.

Joint Working & Code Review

Group Work Notes

Broad approval for joint working initiatives. This is roughly what the best practice guidelines represent (a standardised way to do things), so we anticipate this to be encouraged at a higher level too.

Considerations for joint working:
- Long term support - code and infrastructure
- IP ownership
- Need clear data models
- Explicit requirements
- Portability (docker containers?)

Potential areas/projects for joint working:
- Code review - set of community (best practice?) guidelines and tools to help meet UKAS accreditation and improve quality across the NHS.
- NHS Standardised GATK4 pipeline - software list and parameters.
- Primer database - central repository for sharing primer designs and validation information. Also needs an API to interface with local primer design tools.
- Quality control metrics for sequencing and variant analysis. This would be particularly useful for UKAS accreditation; very little guidance is currently given by UKAS in this area, so community consensus would help drive our adoption of the ISO15189 standard.
- GEL 100k APIs - really well demonstrated by Aled. Identifying the space between GEL and GMCs which is currently in need of tools and development but would affect all the GMCs.
- Standardisation of languages for APIs. The current de facto standard is Python; is this the most appropriate language to be using?

ACTION : Simon Boardman to coordinate an NHS wide audit of GATK pipelines (fastq to vcf). Aim is to identify common software and parameters to create a framework for an NHS Standardised GATK4 pipeline which could be implemented at any centre. By focussing on the software and parameters the resulting standards will be language and hardware agnostic.

GeL data

Too early to agree on large standard schema, agree to work on smaller application but commit to sharing them

Agree to share requriements goals / scope / user stories with group for discussion more widely.

Chris to share template for requirements

Look into project management solutions in github

Data sharing

This is a big topic!

- All agree that data sharing is desirable.
- Lots of potential barriers though.
- Different barriers at different Trusts, e.g. information governance suffers from local policies.
- Differing infrastructures.
- Lots of things that could be done but hard to know how ambitious to be.
- Potential national database on the horizon.

Action

- Recommend all labs must store reported variants in a database and provide minimal dataset (ACGS/Wook).