Providing a research repository for genetic research

David Brown

October 2017

www.genomicsengland.co.uk

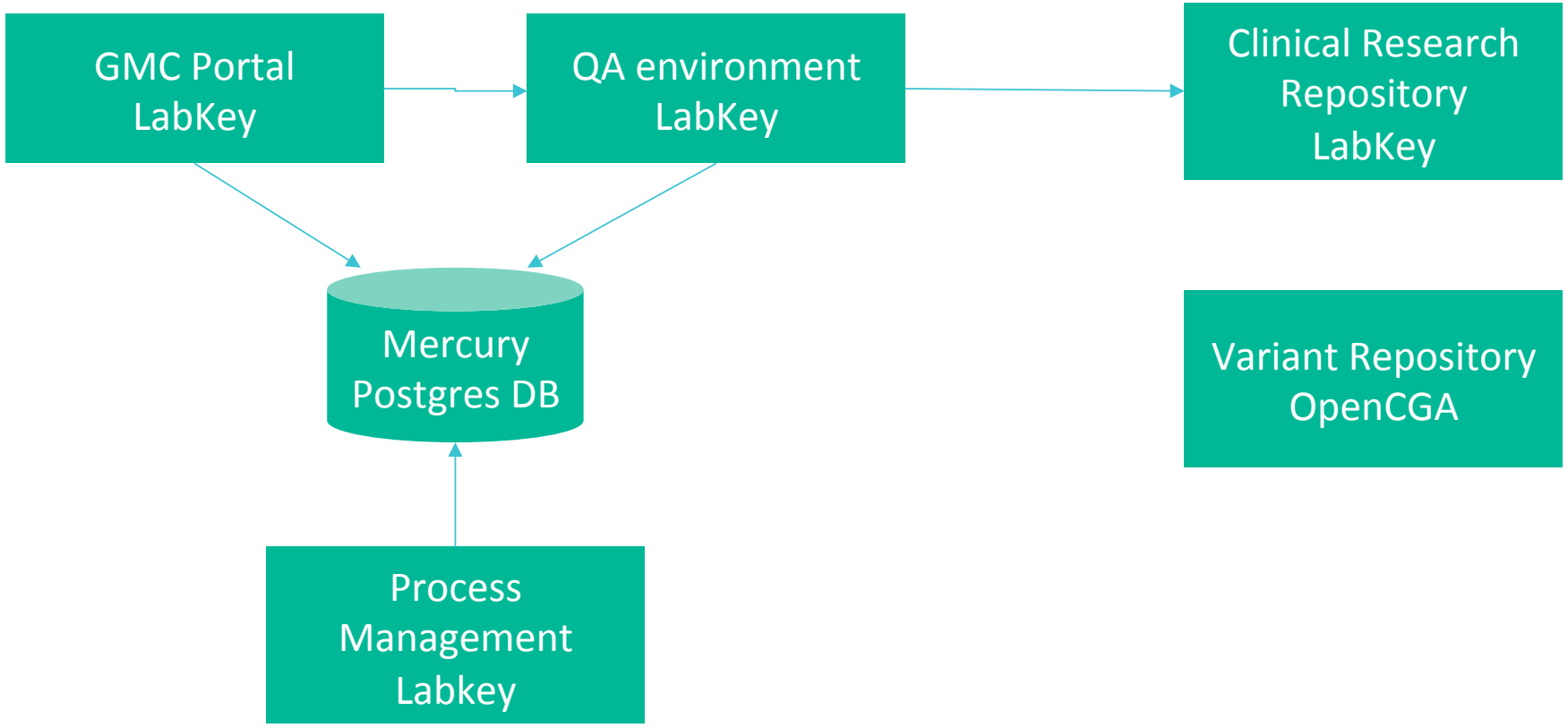
Background to the 100K Genome Project

- To create a research dataset of 100,000 whole genomes collected from patients within the UK National Health Service
- The 100,000 genomes are split between rare disease and cancer
- Rare disease grouped by families (trio's)
- Pave the way for Genetic Medicine in the NHS
  - change in the NHS
  - Exploring processing (centralized v's local)
  - Clinical interpretation – is viable with current technologies and what is needed for it to mature
  - Education in genetic medicine for clinicians in the NHS
  - Drive down costs by working with partners
- Improve the attractiveness of the UK as a place to do research
  - Pharmaceuticals industry
  - Improve access to investigation medicines for NHS patients

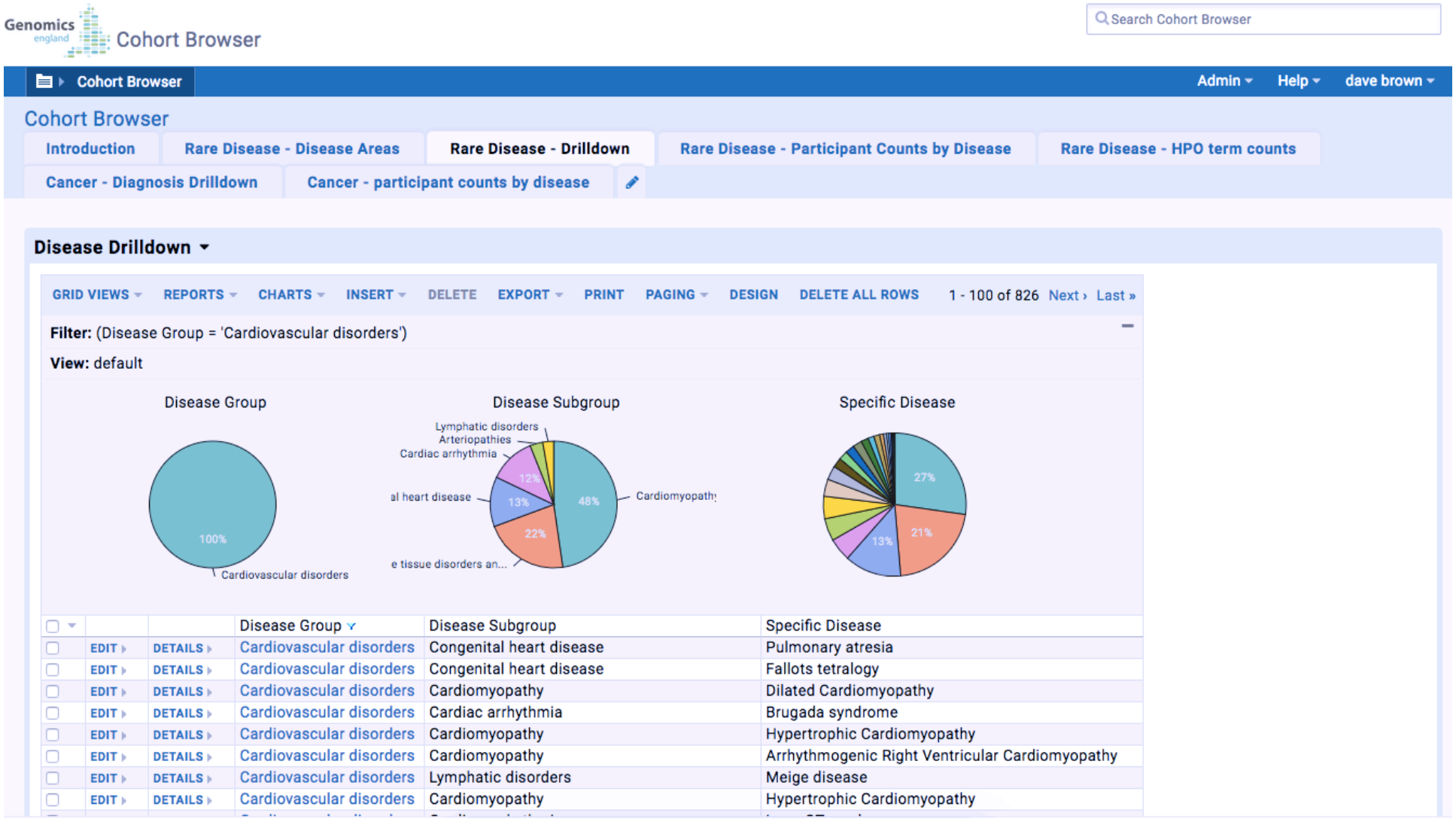
How we currently use LabKey

- To allow physicians to review and amend the participant information they have entered
- Quality assurance and medical review of clinical data
- Production tracking for genomes tracking samples from collection, through to sequence acceptance by Genomics England
- Provision of a Cohort Browser to allow researchers to identify cohorts of participants for their research
- The presentation of datasets for research

Current use of LabKey



Cohort Browser



Current focus – using LabKey to link Clinical and genomic data

- To allow genetic researchers to search for combinations of genetic and phenotypic characteristics to find potentially causal relationships of disease
- Two ‘strands’ of work:
  - What is possible in standard LabKey
  - Integration of LabKey with a variant database (OpenCGA)
- Working closely with the ‘consumers’ of the datasets:
  - Academic research groups
  - The Life sciences industry

Examples of the work so far

- Working with pharmaceutical companies to define the requirements, initially looking into implementing a 2 x 3 matrix

	Homozygous	Heterozygous	Hemizygous
Disease	2,000	3,000	1,500
No Disease	10	100	50

- Working with the Wellcome Centre for Human Genetics to build composite variant calling format files for specific genes from 18,000 participants. Weighting of mutations to identify high confidence mutations.

Holding genetic data in LabKey