2.

PM is about giving the right patient the right treatment at the right time at the right time. Major implications to patient outcomes and costs

Enabled by massive biomedical data sets, omics data, electronic health records ,biobank data.

New generation tools required using efficient algorithms for extracting knowledge hidden within the data

.

This project is about increasing our understanding inter-individual differences in drug disposition,

Strokes are the fourth biggest cause of death and the biggest cause of adult disability cost the country,

There are currently several treatment strategies used on stroke patients not all patients have a good outcome.

3. Drug Metabolism

Evolution –

Process – biotransformation of a xenobiotic compound to enable excretion.

Phase 1 breaking down hydrophobic groups, Phase 2 adding polar groups

DME polymorphisms , include SNPs. INDELs copy number variation.

Pharmacogenomics study affect of genome on drug metabolism

* not a simple link from genome to drug metabolism.-
* a single gene mutation can affect a wide range of metabolomics pathways.
* Complicating factors- environmental and -age , diet, alcohol , drug-drug interactions

Using a complimentary approach - metabolomics

4.Metabolomics

Metabolomics is the endpoint of the central dogma and is directly linked to phenotype, actually represents what is happening in the cell.

Use as orthogonal data to genomics to give new insight eg sickle cell anemia. SNP known for 50years, metabolomics discovered metabolite involved enable a hypothesis about mechanism to be tested.

Annotation a major challenge for untargeted metabolomics. Need for Development of new rigorous and reliable annotation methods tools required to remove this bottle neck.

5.LDA

The goal of LDA represents documents as a mixture of topics.

It backtracks from the words into the documents to find the set of topics that is likely to have generated the document.

This is unsupervised approach so topics not known

The observed variable are the words and the hidden or latent variables are the topic structure and generative process that produced the data.

Documents are the set data for a sample

Words are the individual peaks.

Topics are the unknown metabolites from which the fragments came.

9. My Motivation

Strengths in math and physics but switched from engineering to biophysics as that is where my interests lie.

Msc has been oppurtunty to build update to update knowledge and skills. Learning with access to internet is an amazing experience, suits my learning style.

Over the course of this I have noticed that the bioinformatics methods used are not always the most rigorous available but often chosen by what is known.

Seen it could be hugely beneficial to import methodologies well established in other area eg signal processing (Chip Seq), communications theory microarray data) .But what has really sparked my interest during module in systems biology, is the potential of machine learning. These methods well established by the likes of google. Facebook and amazon. These algorithms have massive potential impact on our ability to extract information and integrate biomedical data .

For me the opportunity to apply Machine learning the in precision medicine where biomedical research is targeted in a clinical setting make a real difference to patients’ lives.