Improved strategies for stratification of disease risk, subtype and therapeutic response in type 2 diabetes and obesity

Supervisors: Professor Mark McCarthy[1,2]
Professor Cecilia Lindgren[3]

Departments: 1. OCDEM, Radcliffe Department of Medicine
2. Wellcome Centre for Human Genetics, Nuffield Department of Medicine
3. Big Data Institute

Background

T2D and obesity are major contributors to global ill health. There is an urgent need for novel strategies for effective prevention and treatment, supported by robust stratification of individual risk and disease subtype that allows increasingly personalised management. However, much of the thinking regarding models for stratification in complex disease is based around notions of discrete, largely homogeneous, subtypes that are more appropriate for rare monogenic diseases and cancers. These models are poorly-suited to the complex, multifactorial aetiology of metabolic diseases such as T2D and obesity: for these, individual risk is influenced by the action of variation at many hundreds (thousands) of mostly common risk variants of small effect, in concert with broadly-shared, pervasive, lifestyle factors.

We recently proposed an alternative conceptual framework better suited to the development and implementation of personalised medicine for complex traits. This **"palette" model** focuses on the intermediate traits that contribute to disease-risk (for T2D: BMI, fat distribution, islet number and function, insulin resistance, islet autoimmunity etc.), each itself subject to multifactorial genetic and environmental influences. Variation in disease-risk, as well as presentation, disease course, and response to specific preventative and/or therapeutic approaches, reflects the individual blend of each of these processes. (By analogy, these processes represent the "base colours" from which can be generated an infinite palette of hues and saturations). This model is consistent with genetic data that point to multiple pathways contributing to disease risk, and the diverse interventions that can ameliorate the diabetic phenotype. *The project proposed here is focused on explicit characterisation of this model using the integration of process-specific genetic risk scores, process-specific biomarkers and "real-world" data to provide an evidence-based framework for personalised prevention and treatment.*

McCarthy and Lindgren are world leaders in the genetics and genomics of T2D and obesity. They lead global consortia that identified and characterised most of the risk loci detected, and, with colleagues, have made substantial strides towards delivering functional insights into disease development. This proposal builds on this platform of discovery by advancing understanding of disease mechanisms in ways that have direct translational relevance.

Hypothesis

This project revolves around the hypotheses that: (a) individual predisposition to T2D and obesity is best understood in terms of the intermediate processes which contribute to disease evolution; (b) process-specific genetic risk scores allow capture of process-specific risk, and will support efforts to achieve more precise phenotypic characterisation through integration with known (e.g. anthropometry, islet cell Ab, fasting insulin) and novel (e.g. omic) biomarkers of disease state; (c) integration of these diverse sources of genetic and genomic information provides cross-sectional and longitudinal descriptions of individual phenotype that will confer clinically actionable information on risk, presentation, disease course, and response to available interventions.

The broad aims of the research to be performed by the fellow are to:

- Assemble process specific genetic risk scores through the integration of large-scale genetic, physiological and genomic data, and explore their relationships to other phenotypes and clinical outcomes;
- Use these data sets to interrogate publicly-available and collaborative proteomic, metabolomics and transcriptomic data, to identify putative biomarkers that capture or mediate process-specific genetic risk;
- Characterise and validate these biomarkers including detailed causal inference and exploration of the impact of environmental risk factors on candidate biomarker levels;
- Generate integrated profiling instruments (genetic + clinical + biomarker) and relate these to disease outcomes (e.g. complications, therapeutic response, disease progression) in longitudinal, prospective data sets.

Description of the work: The primary research focus will be computational and statistical, making use of standard and novel approaches to the interrogation and integration of diverse genetic, genomic and clinical data sets. To achieve the goals described in the previous section, the fellow will be able to access (via supervisors and their collaborative networks, together with publicly-available data) a unique constellation of data and expertise including:

- unparalleled collections of human genetic data describing the relationships between genetic variation and T2D and obesity phenotypes, including: (a) genome wide association (GWAs) and/or exome array data gathered from over 2M individuals (including data from several large biobanks including UK Biobank [500k], GERA [100k], and in the medium-term China Kadoorie Biobank [510K] and Mexico Diabetes Study [200K]); (b) whole genome and exome sequence data from >150K individuals (via the T2D-GENES and TOPMED projects, amongst others);

- rich and complex descriptions of relationships between genetic variation and (a) RNA expression from key tissues relevant to T2D and obesity (islets, fat, muscle, liver, brain, via GTEx and INSPIRE); and (b) circulating 'omic' phenotypes including the plasma proteome, metabolome, and transcriptome (e.g. from Oxford Biobank, DIRECT, UK Twins, Estonian biobank, and other public datasets);

- Prospective follow-up of diabetes outcomes in studies with a variety of genetic, clinical and circulating biomarker data including DIRECT, UKBiobank, Kadoorie and BioVU;

- Extensive personal and collaborative expertise in the analysis of complex genetic, genomic and clinical datasets, including approaches for network generation, causal inference and clinical prediction.

The first phase of these analyses (the first two aims) should be complete by the middle of year 2 of the fellowship. The fellow will be able to pursue follow-up in a variety of alternative directions as outlined in the third and fourth goals. We will be particularly interested in pursuing mechanistic follow-up with colleagues in the new Novo Nordisk Research Centre Oxford. This project will provide the fellow with training in a wide range of computational and statistical techniques, exposure to the biology of T2D and obesity, and powerful opportunities in the emerging field of genomic medicine.

Supervisor's recent relevant publications (McCarthy)

1. Gaulton KJ, Ferreira T, Lee Y, et al. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nature Genetics* 2015;47:1415-25
2. Horikoshi M, Beaumont RN, Day FR, et al. Genome-wide associations for birth weight and correlations with adult disease. *Nature* 2016;538:248-252
3. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. *Nature* 2016;536:41-47
4. Small KS, Todorcevic M, Civelek M, et al. Regulatory Variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. Nature Genetics (in press)
5. McCarthy MI. Painting a new picture of personalised medicine for diabetes. *Diabetologia* 2017;60:793-799

Supervisor's recent relevant publications (Lindgren)

1. Surendran P, Drenos F, Young R, et al. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. *Nat Genet* 2016;48:1151-61
2. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. *Nature* 2016;536:41-47
3. Lessard S, Manning AK, Low-Kam C, et al. Testing the role of predicted gene knockouts in human anthropometric trait variation. *Hum Mol Genet* 2016;25:2082-2092
4. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015;518:187-196
5. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518:197-206