Project Details	
Project Code	MRC21PHBr Lawson
Title	Using population genetics within multi-ethnic Mendelian Randomisation studies
Research Theme	Population Health
Summary	Does drinking moderate amounts of alcohol benefit health - or do only healthy people get to drink moderately? This is just one question solved in epidemiology with Mendelian Randomization - using genetics to separate cause and effect in disease. Genes predispose us to drink more, or less, but most studies are of white European populations. This project will develop statistical tools for drawing causal inference from multi-ethnic biobank comparisons.
Description	Understanding the causal relations between phenotype and disease is key for implementing disease prevention and treatment for population-level health. Mendelian Randomization (MR), that is the use of genetic markers as "instruments" to disambiguate the direction of causality, is a powerful tool to determine whether there is a causal relationship between an exposure (e.g. a risk factor) and an outcome (e.g. a disease) [1]. The use of multiple samples is core to many MR approaches, including 2-sample MR [2] in which the genotype-exposure association is estimated in one sample and the genotype-outcome association estimated in the other. There is an extensive literature in relation to MR including robust methodologies [3] and databases such as MR-Base [4]. MR uses associations between a genetic marker and a trait, but associations vary substantially across populations [5]. Until recently, there have only been substantial high quality data for European populations, meaning that this variation has had little import. Studies such as the China Kadoorie Biobank (www.ckbiobank.org) and the Million Veterans Program (www.mvp.va.gov) containing many African Americans promise to change this rapidly [6]. Traditionally this variation was considered a nuisance for MR, with non-white populations simply ignored in the analysis. We believe it instead provides a new opportunity to finally understand how to fully synthesise data across multi-ethnic groups to strengthen our causal understanding. New methods and assumptions are required to capitalise on diverse genetic resources, but this does not mean such studies will create weaker evidence. We will explore ways in which Genetic diversity can be used as a further robustness check for MR estimates. For example, SNPs that do not vary in effect across populations demonstrate robustness to background correlations, providing added reassurance that they are valid instrumental variables [7]. Beyond robustness to invalid instruments, we can select better instruments using population structure. We rarely know the causal SNPs for each disease as these are in “linkage disequilibrium” with others, i.e. genetically close. Ethnically diverse biobanks can be used to establish "good" instruments, that seem to have the same effect in everyone. To validate the methods, we will develop a simulation of global population history to model the long term change in allele frequencies across populations over time. These simulations will allow testing of MR approaches in the presence of ethnic diversity and allow the design of tests that improve power whilst avoiding bias. [1] Davey-Smith, G. and S. Ebrahim, 'Mendelian randomization': can genetic

	epidemiology contribute to understanding environmental determinants of disease? Int.J Epidemiol, 2003. 32(1): p. 1-22. [2] Pierce, B.L. and S. Burgess, Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. Am J Epidemiol, 2013. 178(7): p. 1177-84. [3] Hemani, G., J. Bowden, and G. Davey Smith, Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum Mol Genet, 2018. 27(R2): p. R195-R208. [4] Hemani, G., et al., The MR-Base platform supports systematic causal inference across the human phenome. Elife, 2018. 7. [5] Lawson, D.J., et al., Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity? Hum Genet, 2019. [6] https://ieureka.blogs.bristol.ac.uk/2020/07/10/most-of-the-world-is-missing-out-on-the-genomics-revolution-why-this-is-bad-for-science/ [7] Bowden, J., G. Davey Smith, and S. Burgess, Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol, 2015. 44(2): p. 512-25.
Supervisory Team	
Lead Supervisor	
Name	Dr Daniel Lawson
Affiliation	Bristol
College/Faculty	Science
Department/School	Mathematics
Email Address	dan.lawson@bristol.ac.uk
Co-Supervisor 1	
Name	Dr Santi Rodriguez
Affiliation	Bristol
College/Faculty	Health Sciences
Department/School	Population Health Sciences
Co-Supervisor 2	
Name	Professor Jack Bowden
Affiliation	Exeter
College/Faculty	College of Medicine and Health
Department/School	Medicine, Nursing and Allied Health Professions
Co-Supervisor 3	
Name	Professor Inês Barroso
Affiliation	Exeter
College/Faculty	College of Medicine and Health
Department/School	Medical School
Co-Supervisor 4	
Name	
Affiliation	
College/Faculty	
Department/School	