**Integrative Epidemiology Unit 4 year PhD studentships:**
**1st year 10 week project outlines**

**1.**

**Title: Development of the "genotype ratio treatment index," a new tool for researching health care delivery and outcomes**

**Supervisors: Professor Ian Day and Professor Frank Windmeijer**

**Background:**
We have shown in a proof-of-principle analysis based on statin prescribing (Day, Davies et al, submitted) that the genotype frequencies in a treatment group can be used to index health care delivery or outcome, where the genotype influences the phenotype for which the prescription is made. For statin prescribing, we hypothesed that *APOE* genotype frequencies in individuals receiving statins should act as a proxy for the prevailing treatment threshold of low density lipoprotein cholesterol (LDLc), since APOE alleles raise LDLc in the order ε2<ε3<ε4 and statin prescribing is at least in part determined by LDLc level. We showed that compared with the general population, and compared with the common, 'wildtype' baseline ε3/ ε3 group, there were almost twice as many ε4 heterozygotes and half as many ε2 heterozygotes. The genotypic ratio alone can be used as an index of treatment; it can also be used, given population data of LDLc distributions for the different genotypes, to backcalculate the average LDLc threshold above which prescribing is taking place in the groups under study (e.g. statin takers; Scottish vs English statin takers; primary versus secondary prevention; socioeconomic classes).

This concept is the focus for a broader development given the diverse range of genetic variants and cardiovascular health care scenarios within which genotype ratio treatment indices may be applicable. In this miniproject, a range of gene loci other than *APOE* influencing LDLc, will be analysed individually, compared, and ways best to combine them, explored.

**Objectives:**
1. To determine genotype ratio treatment indices for 10-15 other independent genetic loci known to specifically influence LDLc levels and relate ratios to LDLc distributions by genotype
2. To compare indices derived from 1 (which can be like risk ratios between population and treatment group; or like odds ratios between genotype groups)
3. To examine genetic models for each locus to determine whether allelic effects are consistent with a multiplicative model, which would enable a reduction of degrees of freedom and composite model of data for all genotypes at a locus. This might also simplify objective 4.
4. To explore approaches to combine the indices derived from 1.

**Methods:**
1. Use existent genetic data and LDLc data in the British Women's Heart and Health Study, and possibly other cohorts
2. Determine relevant genotypic ratios and undertake statistical (parametric likelihood) procedures to achieve objective .
3. To apply appropriate statistical procedures for goals 2, 3 and 4. For goal 2, to draw interpretations with reference to biomedical and functional data bearing on the specificity of effects of each locus examined.

**2.**

**Title: Does prenatal exposure to heavy metals affect IQ and behaviour among children?**

**Supervisors: Sarah Lewis and Carolina Bonilla**

**Background:**
Previous studies have suggested adverse effects of arsenic and lead exposure (even at low levels) on children's cognitive function, including lower IQ scores, impaired attention and memory, and behavioural problems (Lanphear et al., 2005; Rosado et al., 2007). Copper deficiency during gestation can negatively affect motor function, balance and coordination (Georgieff, 2007), whilst excessive free copper may play a role on adult cognitive decline (Salustri et al., 2010). Methylmercury is a known neurotoxicant that can be found as a contaminant in fish. However, it has been difficult to assess its effects on cognitive ability independently of the beneficial effects of fish intake (Daniels et al., 2004).

**Aims:**
To examine the role of prenatal and postnatal levels of metals on the cognitive and behavioural development of children. We will use genetic variants robustly associated with serum concentrations of arsenic, copper, lead and mercury, identified through genomewide association studies (GWAS), as proxies for these exposures and investigate their association with offspring IQ scores and school test results.

Hypotheses: We hypothesise that alleles which increase levels of circulating toxic metals such as arsenic, lead and mercury will be associated with a worse performance in cognitive tests. With respect to copper, both high and low serum levels could potentially lead to an adverse cognitive outcome.

Students will receive training in basic genetic and epidemiological analysis and in Mendelian randomization analysis and interpretation

**References:**
*Lanphear, B. P. et al. (2005). Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. Environmental Health Perspectives 113, 894-899.*

*Rosado, J. L. et al. (2007). Arsenic exposure and cognitive performance in Mexican schoolchildren. Environmental Health Perspectives 115, 1371-5.*

*Georgieff, M. K. (2007). Nutrition and the developing brain: nutrient priorities and measurement. The American Journal of Clinical Nutrition 85, 614S-620S.*

*Salustri, C., Barbati, G., Ghidoni, R., Quintiliani, L., Ciappina, S., Binetti, G., and Squitti, R. (2010). Is cognitive function linked to serum free copper levels? A cohort study in a normal population. Clinical Neurophysiology 121, 502-7.*

*Daniels, J. L., Longnecker, M. P., Rowland, A. S., and Golding, J. (2004). Fish intake during pregnancy and early cognitive development of offspring. Epidemiology 15, 394-402.*

**3.**

**Title: Variations in wound healing loci and their association with injury and scar related outcomes in the ALSPAC cohort.**

**Supervisors: Nic Timpson, Paul Martin and Leila Thuma**

**Background:**
Where adult tissue is damaged, a complex repair process is taking place involving regeneration and acute phase immunological response. Unlike embryonic tissues, adult repair always leads to the formation of a fibrotic scar where the wound has healed, which ultimately can disable proper tissue function [1]. In recent years, research was able to link several genes to the event of scar formation. Knockdown of Ostepontin (OPN) in mice for example resulted in reduced granulation tissue formation and scarring [2]. It also has been indicated that TGF-β1 in conjunction with connective tissue growth factor (CTGF) is promoting scar formation [3]. Most of this data comes from mouse model studies, in humans however, less is known.

Using ALSPAC data work will aim to perform both a candidate driven analysis and a non-hypothesis driven GWAS (the latter being determined by the sample sizes generated from available phenotypic data) comparing individuals involved in an accident developing a scar compared to individuals involved in an accident who did not develop a scar. Where possible, we will attempt to assess differing types of scar tissue and healing related phenotype, however this will again be contingent upon available phenotypic data). ALSPAC is the one of the only cohorts available with data recorded on injury and scarring related outcome alongside comprehensive genetic data on the same participants. The simultaneous availability of these data present a unique opportunity to bring work concerning the biological underpinnings of scarring, fibrosis and inflammation (as developed in Professor Martin's animal models at the UoB School of Biochemistry) forward to testing in human data. We wish to combine specific measures of post injury/accident induced scarring/tissue repair with both genome wide association study data and candidate gene orientated analyses. This will allow us to attempt to establish links between variation in key contributing pathways and real human tissue healing related phenotypes.

We also will extend this research by examining BCG scars as a real time measure of standardised scarring in ALSPAC mothers. BCG injections were routinely given in schools since the 1960s (up until 2005) and lead to a characteristic scar, which individually differs in size. Before injection, a skin test is performed to test for existing Tuberculosis antibodies. If the test is negative, BCG injections are given. In the age group of ALSPAC mothers, up to 70% of school children were immunised each year, making it a great read-out for assessing population wide scarring outcomes. By collecting information about different levels of BCG scarring outcomes in ASLPAC mothers by telephone interviews and combining this data with genome data we want to perform a non-hypothesis driven GWAS. This work has been initiated and will give us the unique opportunity to measure different scarring outcomes population wide to bring work concerning the biological underpinnings of scarring, fibrosis and inflammation (as developed in our animal models at the UoB School of Biochemistry) forward to testing in human data.

**Analysis plan (stage one):**

(i) Assess the depth of data pertinent to scarring phenotypes and establish cases

control series according to differential scarring patterns.

(ii) Perform bioinformatic work up of select genes (LIST) to establish likely functional variants across the coding region and surrounding region.

(iii) Unite both genetic and phenotypic data to undertake tests of association between genetic variation and phenotypic characterisation.

**Analysis plan (stage two):**

(i) Unite genetic and phenotypic data (newly collected BCG scar information) to undertake tests of association between genetic variation and phenotypic characterisation.

(ii) Test selected genes plus surrounding regions (LIST) locally for variant associated with scarring outcome

**Disciplines and training:**
Genomics
Statistical analysis methods
Primary data collection and handling experience
GWAS

**References:**

1. Stramer, B.M., R. Mori, and P. Martin, The inflammation-fibrosis link? A Jekyll and Hyde role for blood cells during wound repair. J Invest Dermatol, 2007. 127(5): p. 1009-17.
2. Mori, R., T.J. Shaw, and P. Martin, Molecular mechanisms linking wound inflammation and fibrosis: knockdown of osteopontin leads to rapid repair and reduced scarring. J Exp Med, 2008. 205(1): p. 43-51.
3. Shi-Wen, X., A. Leask, and D. Abraham, Regulation and function of connective tissue growth factor/CCN2 in tissue repair, scarring and fibrosis. Cytokine Growth Factor Rev, 2008. 19(2): p. 133-44.

**Supervisors/Collaborators**
Dr Nic Timpson, Prof Paul Martin, Leila Thuma

**Title: An investigation into the architecture of eating behaviour: application of the Mandometer to population based data collection.**

**Supervisors: Nic Timpson and Julian Hamilton Shield**

**Background:**
Obesity is extremely common, is a cause of contemporary morbidity, later life risk of disease and is associated with increased mortality. Despite this, levels of obesity are still rising and the factors influencing how people become obese are not fully understood. We are becoming increasingly aware of the potential impact of our environment and our biology on the risk of elevated adiposity, however little is known as to the nature of interactions between these features and how our eating behaviour might influence the likelihood of becoming obese. For instance, work that has been done by our group suggests that it is not just amounts of food and activity that has an influence on the risk of obesity, but the way in which we eat. Results from work targeting the speed with which we eat meals suggests that this and the size of portions we eat influences the risk of becoming obese. The biology of eating behaviour itself, however remains a largely unaddressed field which has implications for public health in this context. Specifically, it is known that speed of eating is related to the development of satiety through regulation of the enteroneural endocrine system[1] and average adiposity[2,3]. Furthermore a recent study which addressed the speed of eating and portion size was shown to be effective in the reduction of body mass index (BMI) when administered in a therapeutic programme[4]. Despite this, there are no recognised methods for the objective assessment of eating speed, nor have such data been considered in light of what might influence variation in this trait in the general population.

Proposed work will build on this information by undertaking an investigation of factors which influence the way in which we eat. Ultimately, we will investigate the biological underpinnings of eating behaviour by examining the way in which people with specific body types and genetic variations consume food. To do this, the primary objective it to use a novel technology ("the mandometer") as a tool for the measurement of eating characteristics within groups of individuals chosen for their levels of obesity or non-obesity and also for the presence or absence of genetic variation known to be associated with adiposity and eating behaviour. To do this we have attained pump priming funding to undertake a project which will assess the mandometer as a tool for population based eating behaviour study. A series of mandometers have been purchased, set up and run in the clinics of the Avon Longitudinal Study of Parents and Children in Bristol where are trained research nurse has collected data on 100 participants so far.

The mandometer (a device used to monitor the rate at which food is consumed by weight, www.mando.se) was developed at the Section of Applied Neuroendocrinology and mandometer Clinic, Karolinska Institute, Stockholm, Sweden. It is a portable weighing scale connected to a small computer that can generate data recording food removal from the plate (which can be viewed graphically, with weight of food (grams) on the y axis and time (minutes) on the x axis). In the therapeutic setting the user puts a measured portion of food determined by a therapist on the scale and the computer records and displays, in real time graphics, the weight loss from the plate as the user eats: time zero on the graph effectively displays total portion size. Removing food from the plate generates a

gradually developing line on a screen that can be compared and matched to a pre-set eating line displaying the speed at which the therapist wants the user to eat.

Alongside these developments in measurement, since 2006 there has been a step change in the methods and results of genetic studies concerning complex phenotypes such as BMI. To this end, studies of common genetic variation (occurring at over 1% in the population) have yielded the first, reliable, genetic associates to adiposity[5,6]. One of the strongest of these effects is that conferred by genotypes at the FTO locus, which has since been termed the "fat mass and obesity related" gene. Relatively little is known about the architecture of this effect although work undertaken suggests that this locus may be conferring effect through and alteration of eating behaviour[6,7,8,10]. With the availability of the "mandometer" instrument, researchers are now presented with a method by which one may simultaneously assess the impact of conscious portion control and monitored consumption rate and also measure the possible impact of heritable variation on patterns of eating behaviour. This proposal will outline a pilot investigation into the use of the mandometer as a tool for measuring eating behaviour in the general population.

The long-term aim of this initiative is to perform a genotype and specific BMI based recall study where the characteristics of eating behaviour are measured and tested according to genotypes across the complete distribution of BMI. To design and prepare for this investigation appropriately, it is critical to discern response rates and gain a good idea of the performance of individuals measured by the mandometer out of a trial setting.

**Work stages:**

Purchase of mandometers and training of research nurse (2 weeks in Sweden). If funded, we will purchase two mandometers and through our collaborative links with The Karolina Institute (who are responsible for the development of the mandometer), organise the training of one research nurse already employed in as part of the Bristol Care of Childhood Obesity Clinic (COCO)[11] (Completed).

Invitation for participation in the study through the ALSAPC study. Through established contact pathways including a mail out to 200 participants (assuming conservative response rates) and a website based volunteer programme, we will invite randomly selected participants to attend the mandometer clinic sessions. Our aim will be to conduct 100 mandometer sessions (Completed).

Running of mandometer clinic sessions is anticipated to occur in the early 2013 (Completed).

Analysis of collected data will be the final stage of this pump priming effort and will occur immediately after the collection of data has been completed in early 2014. Data collected will be assessed for consistency across clinic sessions, compared to an independent collection of mandometer data, described and assessed for response characteristics. The last stage of analysis will be the application of collected data to power calculations which will be used to design the main investigation involving selection of individuals on the basis of already collected genotype and BMI data.

**Disciplines and training**:
Statistical analysis methods

Primary data collection and handling experience
GWAS

**References:**


1. Kokkinos A, le Roux CW, Alexiadou K, et al. Eating Slowly Increases the Postprandial Response of the Anorexigenic Gut Hormones, Peptide YY and Glucagon-Like Peptide-1. J Clin Endocrinol Metab 2009.
2. Llewellyn CH, van Jaarsveld CH, Boniface D, et al. Eating rate is a heritable phenotype related to weight in children. Am J Clin Nutr 2008;88(6):1560-6.
3. Maruyama. K, Sato. S, Ohira. T, et al. The joint impact on being overweight of self-reported behaviours of eating quickly and eating until full: cross sectional study. British Medical Journal 2008;337:a2002.
4. Ford AL, Bergh C, Sodersten P, et al. Treatment of childhood obesity by retraining eating behaviour: randomised controlled trial. Bmj;340:b5388.
5. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316, 889-94 (2007).
6. Scuteri A, Sanna S, Chen W, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genetics 20, e115 (2007).
7. Timpson NJ, Emmett P, Frayling, et al. The FTO/obesity associated locus and dietary intake in children. American Journal of Clinical Nutrition 88, 971-978 (2008).
8. Cecil JE, Tavendale R, Watt P, et al. An obesity-associated FTO gene variant and increased energy intake in children. New England Journal of Medicine 359, 2558-66 (2008).
9. Stratigopoulos G, Padilla SL, LeDuc CA, et al. Regulation of Fto/Ftm gene expression in mice and humans. AJP - Regulatory, Integrative and Comparative Physiology 294, R1185-1196 (2008).
10. Gerken T, Girard CA, Tung YL, et al. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science 318, 1469-1472 (2007).
11. Sabin MA, Ford A, Hunt L, et al. Which factors are associated with a successful outcome in a weight management programme for obese children? J Eval Clin Pract 2007;13(3):364-8.


**Supervisors/Collaborators:**
Dr Nicholas Timpson, Julian Hamilton Shield

**5.**

**Title: Gestational weight gain and daughters' age at menarche: the Avon Longitudinal Study of Parents and Children (ALSPAC)**

**Supervisors: Dr. Abigail Fraser and Prof. Debbie Lawlor**

**Background:** Early life exposures, including in utero, may affect reproductive health later on in the life course. A number of prenatal exposures have been associated with earlier age at menarche,(3) which in turn is associated with a range of adverse metabolic and vascular traits.(2) Hence it is important to identify potential modifiable risk factors for early menarche. Maternal gestational weight gain (GWG) may influence timing of puberty directly by influencing daughters' endocrine system or indirectly through prematurity and both prenatal and post-natal growth.
Two previous studies have yielded conflicting results with the larger study (N=32,218) reporting a U shaped association between GWG and early onset menarche (<11 years). However in this study both exposure and outcome were reported retrospectively.(1)
**Objective:** To examine associations of gestational weight gain and daughters' age at menarche.

**Methods:** Data from the Avon Longitudinal Study of Parents and Children (ALSPAC) will be used. ALSPAC is a prospective population-based birth cohort study that recruited 14,541 pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992 (http://www.alspac.bris.ac.uk.).
The student undertaking this project will:

1. Review the literature to identify other studies that have addressed a similar research objective
2. Identify potential confounders and mediators of the associations of interest
3. Obtain the relevant data from ALSPAC databases
4. Complete necessary data cleaning and deriving of variables
5. Complete analyses
6. Write paper for publication

**Reference List**


1. Boynton-Jarrett R, Rich-Edwards J, Fredman L, et al. Gestational Weight Gain and Daughter's Age at Menarche. J Womens Health (Larchmt ) 2011;
2. Feng Y, Hong X, Wilker E, et al. Effects of age at menarche, reproductive years, and menopause on metabolic risk factors for cardiovascular diseases. Atherosclerosis 2008; 196(2): 590-7
3. Maisonet M, Christensen KY, Rubin C, et al. Role of Prenatal Characteristics and Early Growth on Pubertal Attainment of British Girls. Pediatrics 2010; 126(3): e591-e600

**Title: Diversity in methylation profiles from multiple tissue types from individuals of the Cleft Collective Cohort studies**

**Supervisors: Beate St. Pourcain and Caroline Relton**

**Background:**
DNA methylation is an epigenetic mechanism, which leads to heritable phenotypic alterations within the absence of DNA sequence changes and plays a key role in gene regulation and development[1]. On a molecular level, the methylation of DNA occurs at cytosines of CpG sequences, which is preserved during replication in vertebrates, and can thus be mitotically inherited. Methylation of CpG clusters or CpG islands within gene promoters can silence gene expression as DNA methylation is generally linked with chromatin condensation via histone modifications[2]. It has been shown that mammalian somatic tissues are differentially methylated at unique sequence sites within the genome (acting like a fingerprint) [2,3] and that these differentially methylated regions are evolutionary conserved[3].

This project will investigate the methylation profile in cells and tissues collected from patients with cleft lip and/or palate (CL/P) and their parents to identify overarching similarities and dissimilarities in methylation patterns from different cells and tissues. CL/P is a common congenital disorder involving a facial cleft structure, which affects about a thousand children born in the United Kingdom (UK) each year. The cleft structure in children will be repaired through a surgical intervention during the first year of their lives (lip operation at 3 months of age, palate operation at 9 months of age). This offers the opportunity to collect less accessible human tissue types, which are a by-product during surgery. The project will provide unique insights into the potential role of DNA methylation in CL/P.

**Objectives:**
1. To identify similarity and diversity between methylation profiles of biological samples in children with a CL/P (palate or lip tissue, saliva, venous blood, and fibroblast cell lines) and their parents (saliva, venous blood)
2. To identify similarities in methylation profiles within families
3. To interrogate the potential functional relevance of differentially methylated regions identified in the tissues studied using a range of bioinformatic tools

**Methods:**
1. We will perform basic statistical analyses between the methylation profiles derived from saliva and blood in parents and children using mixed linear models
2. We will also explore the correlation between methylation profiles from blood, saliva, fibroblasts and palate/lip tissues in children only using hierarchical clustering approaches and other analysis approaches which will be dictated by the data generated.
3. Systematic search of publicly available databases for information on functional relevance of differential methylation sites observed.

**References:**

1. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204–220 (2013).

2. Ohgane, J., Yagi, S. & Shiota, K. Epigenetics: the DNA methylation profile of tissue-dependent and differentially methylated regions in cells. *Placenta* **29 Suppl A**, S29–35 (2008).
3. Nagase, H. & Ghosh, S. Epigenetics: differential DNA methylation in mammalian somatic tissues. *FEBS J.* **275**, 1617–1623 (2008).

**7.**

**Title: Using a Mendelian randomization framework to determine whether there is a causal association between early menarche and depression**

**Supervisors: Sarah Lewis, Carol Joinson, George Davey Smith**

**Background:**
The project seeks to determine whether early pubertal timing leads to an increased risk for depression in adolescent girls and boys and in women. Observational studies have found that depression is more common in young women who have undergone a relatively early menarche. Inferring causation is often a problem with epidemiological data because it is not always clear whether the observed association between an exposure and outcome variable is due to common effects of a third unobserved variable (confounder). Potential confounders were rarely considered in previous studies reporting an association between early menarche and depression. For instance, exposure to socioeconomic disadvantage and father absence are associated with earlier timing of menarche and with increased risk of depression in teenage girls (Joinson et al. 2011). Less work has been done on the relationship between age of puberty and depression in men, but age of puberty in men is more difficult to define.

Mendelian randomization (MR) has been proposed as a method to examine the causal relationship between an exposure and an outcome when confounding is believed to be likely (Davey Smith and Ebrahim, 2003). MR exploits Mendel's second law, that the inheritance of one trait is independent of (randomized with respect to) the inheritance of other traits, and more importantly, independent of lifestyle and other confounding factors. MR uses a genetic marker that has a *known* relationship with the exposure as an *instrumental variable*, in an analogous way to randomization in a clinical trial, to estimate the causal relationship between exposure and outcome. If there is evidence that a gene is closely linked to the exposure, it can be assumed that the gene is not itself associated with any confounders. Genetic variation associated with the exposure is, therefore, used as a *non-confounded* proxy for that exposure. A common polymorphism in the LIN28B gene (rs314276) was found in an early genome wide association study to be associated with early timing of menarche (Ong et al. 2009). More recently many more such genetic variants have been identified. We wish to use MR techniques to provide a *un-confounded* assessment of the relationship between early menarche and depression in the ALSPAC cohort, which will allow us to determine the direction of effect between the two.

**References**

1. Joinson C, Heron J, Lewis G, Croudace T, Araya, R. Timing of menarche and depressive symptoms in adolescent girls from a UK cohort. Br J Psychiatry 2011;198: 17-23.
2. Davey Smith G and Ebrahim S, 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003 Feb;32:1-22
3. Ong KK et al. Genetic variation in LIN28B is associated with the timing of puberty. Nat Genet 2009;41(6):729-33.
4. Palmer TM et al. Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. Am J Epidemiol 2011 Jun 15;173(12):1392-403.

**8.**

**Title: A systematic review of mechanistic studies of calcium and prostate cancer**

**Supervisors: Sarah Lewis, Mike Gardner and Richard Martin**

**Background:**

Many laboratory experiments are performed to inform human health. These types of mechanistic studies compliment epidemiological findings and can offer insights into biological plausibility and pathways between exposure and disease. In 1965 Bradford Hill stipulated that biological plausibility was one criteria on which a causal link between exposure and disease could be judged. However, given the plethora of published mechanistic studies, some way of judging the quality and reproducibility of these studies is needed. Experimental studies have been criticized recently due to the poor quality of many studies which means they are not reproducible. Particular problems which have been highlighted are low statistical power incorrect statistical analysis, lack of blinding and randomization, lack of quality-control mechanisms, deficiencies in reporting, and publication bias. With this in mind it is therefore important to identify findings which are consistent across studies and which may advance scientific knowledge. We have been developing a protocol which can be used to systematically review mechanistic studies in order to be able to reach some conclusions on the quantity and quality of evidence which supports a particularly hypotheses. In this project we would like the student to carry-out literature searches to identify all mechanistic studies which link calcium to prostate cancer. The student will then synthesis this evidence, and present it in a way which means that conclusions can be drawn about the likely mechanisms explaining the observed associations and the quality of evidence which underpins these.

Students will receive training in literature searching, data extraction, assessing study quality and evidence synthesis.

**9.**

**Title: Deriving a measure of maturation status based on percentage of predicted adult stature.**

**Supervisor: Prof Kate Tilling, Dr Laura Howe, Dr Sean Cumming (University of Bath)**

**Background:**

Pubertal status is an important determinant of many health-related behaviours and outcomes. It is therefore an exposure of interest for numerous outcomes, as well as being a confounder in many exposure-outcome associations. Standard methods to assess pubertal status are sometimes seen as an invasion of personal privacy, and therefore response rates from research participants tend to be low. There are also questions over the validity of methods that rely on self (or parent) reported pubertal status. A non-invasive method of assessing maturation status based on the percentage of adult stature (predicted from parental height) that a child has attained has been developed.(1, 2) The percentage of predicted final height for an individual can be compared with age- and sex-specific references to give z-scores, which can then be used to assess the degree to which a child is advanced or delayed in maturation status. The aim of this project is to apply this method to data from participants of the Avon Longitudinal Study of Parents and Children (ALSPAC). Once the z-scores have been derived, you will write documentation to guide other researchers in the use of these variables. You will also compare these measures with other puberty assessments: age of menarche for girls, and for boys and girls, assessments of puberty based on questionnaires in which the child's parent or guardian (most often mother) was asked to identify which line-drawing most closely resembled the child. This method(3, 4) describes 'Tanner'stages ranging between I and V, with I being least sexually mature. Since heights at age 18 are also now available for the ALSPAC cohort, we can compare the predicted final heights with the actual measured heights at age 18.

**Objectives:**

1. To develop z-scores of maturation status for all participants in ALSPAC who have data on height, weight, and parental self-reported heights
2. To compare this non-invasive assessment of maturation with other pubertal measures at different ages
3. To compare predicted final heights with actual measured heights at age 18

**Methods:**

1. Obtain the relevant data from ALSPAC databases
2. Complete data cleaning and derive z-scores of maturation status
3. Compare z-scores of maturation status with other puberty measures
4. Compare predicted final heights with measured heights at age 18
5. Write documentation to guide other researchers in the use of derived variables
6. Write paper for publication

**References:**


1. Khamis HJ, Roche AF. Predicting adult stature without using skeletal age: the Khamis-Roche method. Pediatrics 1994;94:504-7.
2. Malina RM, Dompier TP, Powell JW et al. Validation of a noninvasive maturity estimate relative to skeletal age in youth football players. Clin J Sport Med 2007;17:362-8.

3. Morris NM, Udry JR. Validation of a self-administered instrument to assess stage of adolescent development. Journal of Youth and Adolescence 1980;9:271-80.
4. Duke PM, Litt IF, Gross RT. Adolescents' self-assessment of sexual maturation. Pediatrics 1980;66:918-20.

**10.**

**Title: Y chromosome and mitochondrial DNA to inform genetic association studies: *FTO* and body weight as an example**

**Supervisors: Dr Santi Rodriguez, Dr Nic Timpson and Dr David Evans**

**Background:**
Both Y chromosome and mitochondrial DNA (mtDNA) differ in relation to autosomal loci in many respects, including their inheritance pattern. The nonrecombining region of the Y chromosome is paternally inherited, with no recombination with its homologous chromosome pair (X chromosome). The mtDNA is maternally inherited. Genetic variation in the Y chromosome and mtDNA define groups of haplotypes which share common ancestor with a SNP mutation. These groups of haplotypes, or haplogroups, have been widely characterised in worldwide populations. Phylogenetic relationships between haplogroups are well known and there is a good knowledge of the frequency and types of haplogroups present in all geographical regions throughout the world. For example, haplogroup R1b is very frequent in Europe and infrequent or absent in the other continents. There is detailed information about around thirty subclades of haplogroup R1b, including the specific SNPs that define them, approximate time of origin, place of highest frequency and most prevalent ancient ethnic group present in different European sub-regions. Therefore, both Y and mtDNA haplogroups can be used to identify groups of individuals within major ethnic groups based on their non-recombining ancestry.

Our hypothesis is that Y chromosome and mtDNA haplogroups can be used to inform association studies in large epidemiological cohorts. The consideration of homogeneous groups of individuals according to non-recombining ancestry can identify the haplogroup/s accounting for the association in genetic association studies. This will have implications to the interpretation of genetic associations. Our aim is to refine the reported association between *FTO* and body weight by analysing subgroups of individuals grouped according to their non-recombining ancestry.

**Objectives:**
1. To test the association between *FTO* and body weight related traits in groups of individuals with the same haplogroup.
2. To compare the associations found in this project with previously reported associations using the whole sample of Europeans in ALSPAC.

**Methods:**
1. We will analyse the haplogroups already available for 20,000 children and mothers in ALSPAC, generated by 23 and me.
2. We will analyse phenotypic information for body weight traits from ALSPAC.
3. We will perform basic statistical analyses for the genetic association studies, including linear and multiple regression between *FTO* genotypes and body weight related traits.

**11.**

**Title: Investigating Genetic Interactions Between Mitochondrial DNA and Autosomal Loci**

**Supervisors: Dr Santi Rodriguez, Dr Nic Timpson, Dr David Evans,**

**Background:**
Mitochondrial DNA (mtDNA) differs in relation to autosomal loci in many respects, including its pattern of maternal inheritance. Genetic variation in mtDNA defines groups of haplotypes which share a common ancestor with a SNP mutation. These groups of haplotypes, or "haplogroups", have been widely characterised in worldwide populations and provide useful information on the maternal ancestry of an individual. Our hypothesis is that mtDNA variants and haplogroups interact with variants on the autosome to produce phenotypic variation and hence explain some of the "missing heritability" observed for many common traits and diseases. As mitochondria are intimately involved in aerobic respiration, obesity related traits seem like a logical place to start the search for interacting variants.

**Objectives:**
1. To investigate the existence of statistical interaction between mitochondrial SNPs/ haplogroups and autosomal loci in determining childhood obesity.

**Methods:**
1. Individuals will be classified according to mitochondrial SNPs and/or Haplogroup.
2. We will analyse phenotypic information for body weight traits from ALSPAC.
3. Statistical models will be fit to the data that model the relationship between autosomal SNP, mitochondrial SNP/haplogroup and obesity.

**12.**

**Title: Fetal and postnatal growth and the modifying effect of breastfeeding on long-term levels of circulating insulin-like growth factor: analysis of 17,000 children in the PROBIT trial**

**Supervisors: Kate Tilling and Richard Martin**

## Background:

Insulin-like growth factor (IGF)-I is a major regulator of growth and is associated with cancer and CVD outcomes in adulthood. In the large PROBIT study of over 17,000 children, this project will compare children born SGA (birth weight <10th percentile), those born large for gestational age (LGA, >90th percentile for gestational age), and the remaining born appropriate for gestational age (AGA). Multilevel statistical models will be used to account for clustered measurement and to control for maternal and paternal height and BMI, maternal education, geographic region, urban vs rural residence, and the child's exact age at the follow-up visit. It also examines the long-term effects on circulating IGF-I of catch-up growth among SGA infants (>0.67 SD increase in weight-for-age) and catch-down growth among LGA infants (>0.67 SD decrease in weight-for-age) in the first 3-6 postnatal months. If there are any main effects, then you would be to determine if breastfeeding [either as randomized, i.e., in intention-to-treat analyses, or as fed (exclusive breastfeeding for 3 months, defined as yes or no] interact with SGA or LGA birth with respect to the IGF outcomes at 11.5 years of age.

**13.**

**Title: Are differences in methylation in cord blood DNA associated with prenatal exposure to alcohol?**

**Supervisors: Dr Luisa Zuccolo and Prof Caroline Relton**

**Background:**
The mechanisms that link intrauterine alcohol exposure to neuro-developmental consequences in the fetus are largely unknown, but they could include developmental programming effects, such as epigenetic changes caused by ethanol crossing the placenta. The aim of this mini project will be to investigate cord-blood DNA methylation profiles of offspring differentially exposed to alcohol in utero, in candidate genes in pathways likely to be targeted/disrupted by early ethanol exposure.
The study will be based on the Avon Longitudinal Study of Parents and Children (ALSPAC). Methylation analyses have already been carried out, interrogating 96 CpG sites within systematically prioritised candidate genes. To limit confounding and bias, which could arise comparing offspring whose mothers reported drinking in pregnancy to those whose mothers reported abstaining, the project will compare DNA methylation in offspring grouped according to maternal genotype. A functional SNP in the alcohol-metabolising gene ADH1B is known to affect the mothers' physiological response to alcohol and their propensity to drink, but is not related to other socio-economic and behavioural factors that often confound intrauterine alcohol-offspring outcome associations. This genetic variant will be used as proxy for maternal drinking. This is one of the first times that such an approach, known as Mendelian Randomization, has been used to study determinants of DNA methylation. Further analyses based on data from an epigenome-wide association study will assess the specificity of methylation signatures in this panel of candidate genes compared to the rest of the genome.
This project has huge potential for publication in a high impact scientific journal, given the importance of understanding the mechanisms linking prenatal alcohol exposure to offspring outcomes.

**Objective:** To examine the association of prenatal alcohol exposure with methylation of cord blood DNA in the Avon Longitudinal Study of Parents and Children (ALSPAC).

**Methods:**
Review the literature on prenatal alcohol exposure and changes in offspring DNA methylation
Obtain the relevant data from supervisors
Complete analysis of potential differentially methylated regions in candidate CpG (VeraCode data)
Compare signals to those in other CpGs outside the panel of candidate genes (450k data)
(optional) Validate methylation signals through pyrosequencing
Write up for publication

**14.**

**Title: Neurodevelopmental consequences of prenatal alcohol exposure: a population based MRI study**

*(Note: this topic includes at least two PhD and appointed students will be expected to define their specific area of research within this broad area )*

**Supervisors: Dr Luisa Zuccolo and Prof Caroline Relton**

**Background:**
There is clear evidence that heavy alcohol intake during pregnancy disrupts normal brain development with long-term adverse cognitive and behavioural consequences in offspring [1,2]. While gestational alcohol *abuse* remains a problem, occasional use of *small-to-moderate amounts* of alcohol is a more prevalent behaviour during pregnancy, and therefore potentially associated with greater population-level harms. However, it is unknown whether this drinking pattern is harmful for the developing brain, since most recent epidemiological studies fail to find adverse effects [3]. Nevertheless, evidence is emerging from different study designs more robust to biases including confounding, suggesting that even small amounts of alcohol during pregnancy could potentially affect neurodevelopment and result in lower cognitive and academic scores [4,5], as well as from animal models [6]. Epigenetic marks (e.g. DNA methylation changes) resulting in altered regulation of gene expression provide one likely mechanism for the neurodevelopmental effects of alcohol, with key developmental stages of intrauterine and early life thought to be sensitive periods when the epigenome is particularly responsive to external stimuli such as alcohol [7], and preliminary evidence linking epigenetic marks to future disease risk. To our knowledge, the subtle effects of alcohol use on the brain have not been studied yet in large-scale population samples, nor has the potential mediating role of DNA methylation on such effects.

**Objectives:**

The overall aim of this project is to explore the associations of alcohol consumption during pregnancy with offspring structural brain morphology (structural properties of white and grey matter), with particular emphasis on their causal nature and on the role played by DNA methylation. Specific objectives are:

1. To study the effects of prenatal exposure to moderate levels of alcohol on structural brain morphology in school age children and adolescents.

2. To explore the causal nature of the observed associations using causal analysis methods such as cross-cohort comparisons, negative control and Mendelian randomization methods.

3. To investigate whether DNA methylation is a causal mediator of the above associations, by studying whether prenatal alcohol exposure is causally associated with DNA methylation changes detected at birth and persisting into adolescence, and whether such alcohol-induced methylation changes are causally associated with brain MRI and DTI outcomes.

**Methods:**

There are a number of potential PhD projects that could be undertaken based on this proposal. The student will be encouraged to develop a work plan based around their interests.

All potential studentships will involve the use of multiple data sources including three ongoing cohort studies: Generation R [9], the Avon Longitudinal Study of Parents and Children (ALSPAC) [10] and the Saguenay youth study [11], and a number of different methodologies. Exposure and outcome measures can be readily harmonized across the 3 studies, and a wealth of data exists already to allow investigation of possible confounding and mediating effects (e.g. prenatal and postnatal environmental and psycho-cognitive assessments, genome-wide genetic scans and DNA methylation data). Together, these 3 studies provide enough power to detect or exclude even subtle effects (e.g. [8]), and enough design differences and a wealth of harmonised data to rule out major biases and confounding as explanation of the results. Epigenome-wide DNA methylation data available in all 3 cohorts will allow the interrogation of the mediating role of DNA methylation by using two-step Mendelian randomization, a newly proposed method for establishing whether DNA methylation causally explains (at least part of) a known exposure-outcome association [14]. Uncovering a potential causal pathway will add strength to the accumulating evidence for long-term neurodevelopmental consequences of prenatal alcohol exposure.

The student(s) will have the opportunity to familiarise with different research methodologies. These will include epigenomics and neuroimaging data processing and analyses (e.g. see http://www.godmc.org.uk/, http://www.ariesepigenomics.org.uk and [12]), and causal analysis methods such as cross-cohort comparisons, negative control methods, and Mendelian randomization [13,14].

**Disciplines, research skills and training**

Epidemiology
Causal analysis methods
Epigenetics
Neuroimaging
The candidate will be provided with opportunities for international collaborations (Rotterdam or Toronto)

**References**

1. Irner T.B., Substance exposure in utero and developmental consequences in adolescence: a systematic review. Child Neuropsychol, 2012. 18(6): p. 521-49.
2. Lebel C., et al., A Longitudinal Study of the Long-Term Consequences of Drinking during Pregnancy: Heavy In Utero Alcohol Exposure Disrupts the Normal Processes of Brain Development. The Journal of Neuroscience, 2012, 32(44): 15243-15251.
3. Flak A.L., et al., The Association of Mild, Moderate, and Binge Prenatal Alcohol Exposure and Child Neuropsychological Outcomes: A Meta-Analysis. Alcohol Clin Exp Res, 2013 Aug 1. [Epub ahead of print].

4. Zuccolo L., et al., Prenatal alcohol exposure and offspring cognition and school performance. A 'Mendelian randomization' natural experiment. Int J Epidemiol, 2013. 42(5):1358-70.
5. Lewis S.J., et al., Fetal alcohol exposure and IQ at age 8: evidence from a population-based birth-cohort study. Plos One, 2012 2012;7(11):e49407.
6. Valenzuela C.F., et al., Does moderate drinking harm the fetal brain? Insights from animal models. Trends in neurosciences, 2012;35(5):284-292.
7. Haycock P.C., Fetal alcohol spectrum disorders: the epigenetic perspective. Biol Reprod, 2009 Oct;81(4):607-17.
8. El Marroun H., et al., Prenatal tobacco exposure and Brain morphology: A prospective study in young children. Neuropsychopharmacology, 2013 Oct 7. [Epub ahead of print].
9. Jaddoe V.W., et al., The Generation R Study: Design and cohort update 2012. Eur J Epidemiol, 2012 Sep;27(9):739-56.
10. Boyd A., et al., Cohort Profile: The 'Children of the 90s' - the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol, 2013. 42(1): p. 111-27.
11. Pausova Z., et al., Genes, maternal smoking and the offspring brain and body during adolescence: design of the Saguenay youth study. Hum Brain Mapp, 2007;28:502-18.
12. White T., et al., Pediatric population-based neuroimaging and the Generation R Study: the intersection of developmental neuroscience and epidemiology. Eur J Epidemiol, 2013. 28(1): p. 99-111.
13. Gage, S.H., et al., Using Mendelian randomisation to infer causality in depression and anxiety research. Depress Anxiety, 2013 Jul 11. [Epub ahead of print].
14. Relton C.L. and Davey Smith G., Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. Int J Epidemiol, 2012;41(1):161-176.

**Title: Alcohol and prostate cancer risk – Mendelian randomization in the largest case-control study to date.**

**Supervisors: Dr Luisa Zuccolo, Dr Neil Davies**

**Background:**
Alcohol is a carcinogen and understanding whether drinking alcohol causes prostate cancer may provide insights into its aetiology and potential for prevention. Prospective studies have shown an increased prostate cancer risk among heavy drinkers (1,2), however two earlier systematic reviews could not provide a robust answer to the role alcohol plays in prostate cancer (3,4), possibly because of methodological limitations including confounding, reverse causation, selection and detection bias, exposure misclassification and low statistical power. To overcome some of these problems we will use an approached known as "Mendelian randomization". This uses common genetic variants which affect the propensity to drink and/or the physiological response to alcohol but are not related to other socio-economic and behavioural factors that often confound alcohol-prostate cancer associations, as proxies for self-reported alcohol consumption. Thus we will use these variants to accurately estimate the effects of drinking alcohol.
This project has huge potential for publication in a high impact scientific journal, we have a very large sample and this question is very important in terms of improving our understanding of alcohol-related harm and ultimately helping find ways to prevent prostate cancer.

**Aim:**
To obtain estimates of the causal (unconfounded and unbiased) effects of alcohol on prostate cancer incidence and progression, by investigating the distribution of common genetic variation in alcohol metabolism genes among 22,000 cases and 22,000 controls from a multi-centre international consortium: PRACTICAL- *Prostate Cancer* Association Group to Investigate Cancer Associated Alterations in the Genome.

**Objectives:**
7. Identify and review recent studies on alcohol and prostate cancer.
8. Complete necessary data cleaning and deriving of variables.
9. Investigate the association of the genotype with potential confounding variables.
10. Investigate the association of the genotype with prostate cancer.
11. Write and publish a research paper to disseminate the findings.

**Skills to be gained:**
- Working as a member of a team to accurately clean, construct and analyse datasets using genetic variants.
- Understanding of candidate gene association studies and Mendelian randomization.
- Experience preparing papers for submission and peer review.

**References:**
(1) Zuccolo L, Lewis SJ, Donovan JL, Hamdy FC, Neal DE, Davey Smith G. Alcohol consumption and PSA-detected prostate cancer risk-A case-control nested in the ProtecT study. Int J Cancer. 2013 May 1;132(9):2176-85.

(2) Gong ZH, Kristal AR, Schenk JM, et al. Alcohol consumption, finasteride, and prostate cancer risk results from the prostate cancer prevention trial. Cancer 2009; 115: 3661–9.
(3) Middleton FK, Chikritzhs T, Stockwell T, et al. Alcohol use and prostate cancer: a meta-analysis. Mol Nutr Food Res 2009; 53: 240–55.
(4) World Cancer Research Fund and American Institute for Cancer Research. Systematic literature review on diet and prostate cancer: alcoholic beverages. In: Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Second Expert Report. Washington DC: AICR, 2007. 397–447.

**16**

**Title: A genome-wide scan of alcohol sensitivity in teenagers**

**Supervisor: Dr Luisa Zuccolo**

**Background:**

To date, very few common genetic variants are known to be associated with alcohol behaviour in populations of European ancestry. In particular, the genetic component of dimensions of teenage drinking has not been explored yet, despite suggestions that alcohol behaviour has a strong heritable component even at a young age.

Alcohol sensitivity is one of the heritable factors conferring susceptibility to risk of alcoholism in adults and problem drinking in teenagers. Alcohol sensitivity has been shown to correlate well with items from the self-rating of the effect of alcohol (SRE) questionnaire, especially when the questionnaire is administered close in time to the beginning of one's "drinking career". To date, no studies have been conducted to search for the common genetic variants underpinning alcohol sensitivity as measured by the SRE scale. Such genetic variants would be promising candidates for explaining some of the variability in alcohol behaviour at the population level, particularly in terms of the susceptibility to problem drinking.

**Objective:**

To determine whether genome-wide common genetic variation influences alcohol sensitivity in adolescents, measured through the self-rating of the effect of alcohol scale in the Avon Longitudinal Study of Parents and Children (ALSPAC).

Detailed information on level of response to alcohol consumption throughout the teenage years is available, so that both cross-sectional and longitudinal analyses can be performed.

**Methods:**

1. Review the literature on genetic variation and alcohol sensitivity
2. Identify relevant alcohol phenotypes and agree an analysis plan with supervisors
3. Obtain the relevant data from ALSPAC databases/supervisor
4. Complete statistical analysis
5. Investigate feasibility for replication of main findings in an independent sample
6. Write up for publication

**References:**

(1) Schumann G, et al. (2011) Genome-wide association and genetic functional studies
identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. Proc Natl Acad Sci USA 108(17):7119–7124.

(2) Schuckit MA, Smith TL, Heron J, Hickman M, Macleod J, Lewis G, Davis JM, Hibbeln JR, Brown S, Zuccolo L, Miller LL, Davey-Smith G. Testing a level of response to alcohol-based model of heavy drinking and alcohol problems in 1,905 17-year-olds. Alcohol Clin Exp Res. 2011 Oct;35(10):1897-904.

**Title: Alcohol consumption in pregnancy and pregnancy complication: the Avon Longitudinal Study of Parents and Children**

**Supervisors: Dr Luisa Zuccolo and Dr Abigail Fraser**

**Background:**
It has been hypothesized that alcohol may play a role in placenta-associated syndromes though the exact mechanisms remain unclear. Numerous studies have examined associations of maternal alcohol intake in pregnancy with birthweight, intra uterine growth restriction, preterm and still birth. Results of these studies are inconsistent which may be due to differences in study design, exposure definition, and degree of adjustment for potential confounders. However, to the best of our knowledge the associations between maternal alcohol intake with the two most common pregnancy complications namely hypertensive disorders of pregnancy (HDP, including gestational hypertension and pre-eclampsia) and gestational diabetes have not been examined.

**Objective:**
To examine the association of maternal alcohol intake in pregnancy with pregnancy complications in mothers of the Avon Longitudinal Study of Parents and Children (ALSPAC).

Detailed information on alcohol consumption before and for each trimester in pregnancy is available.

**Methods:**
1. Review the literature on alcohol intake in pregnancy and pregnancy complication as well as other relevant outcomes that could inform this work
2. Define relevant categories of exposure
3. Identify potential confounders and agree an analysis plan with supervisors
4. Obtain the relevant data from ALSPAC databases/supervisors
5. Complete derivation of variables and analysis
6. Write up for publication

**References:**

Salihu, H. M., et al. (2011). "Impact of prenatal alcohol consumption on placenta-associated syndromes." Alcohol 45(1): 73-79.

Henderson, J., et al. (2007). "Systematic review of effects of low–moderate prenatal alcohol exposure on pregnancy outcome." BJOG: An International Journal of Obstetrics & Gynaecology 114(3): 243-252.

**Title: Approaches to data integration and visualisation in multi-omics and epidemiological data**

**Supervisors: Tom Gaunt, Colin Campbell, Jose Soeane**

**Background:**
New technologies mean we are now able to collect data on thousands or millions of different molecular variables in parallel in an individual sample. The application of these in large population samples like the Avon Longitudinal Study of Parents and Childhood (ALSPAC) is creating extremely powerful data resources for understanding the mechanisms of health and disease. In addition, the availability of a wealth of molecular data in public databases enables rich annotation and enhancement of population-based studies. However, data integration and analysis of these high-dimensional datasets presents specific computational challenges, and new methods are required to maximize their potential.

**Objectives:**
The aim of this bioinformatics project is to develop new approaches to data integration and visualisation in high-dimensional omics datasets.

**Plan:**
Beginning with methylation, genetic, expression and metabolomic data generated on samples from the ALSPAC cohort, the student will develop methods to integrate these datasets, identify interesting patterns in the data and relate these to data from public databases. The project will initially be focused on one specific mechanism or phenotype (to be agreed with student), but the expectation would be that the methods developed will be widely applicable.

**Proposed methods/technologies:**
Methods will be developed as part of the project, but amongst many others may involve kernel methods, graph-based data integration, GPGPU approaches, data federation and consolidation

**19**

**Title: Aggregating methods for analysis of association between methylation and phenotypes or outcomes**

**Supervisors: Tom Gaunt, Colin Campbell, Jose Soeane**

**Background:**
DNA methylation plays an important role in the regulation of genes. We have measured DNA methylation in the peripheral blood of 1000 participants in the Avon Longitudinal Study of Parents and Children (ALSPAC) at three timepoints in childhood. Whilst single-point (i.e single CpG site) association analyses can be informative, they result in a large number of tests and make no use of prior information about the functional importance of the genes in which these sites are based. Aggregating methods which combine information across multiple CpG sites based on genome annotations and pathway information have the potential to reduce the number of tests and increase the power.

**Objectives:**
The aim of this bioinformatics project is to utilize pathway information to reduce the dimensionality and improve the power of epigenetic association studies.

**Plan:**
The project will utilize correlation and pathway data to construct defined sets of CpG sites for a specific research question (to be agreed with student). These sets will then be tested for association with the agreed phenotype. A post-hoc analysis will be performed to determine whether a single-point analysis would have been sufficient.

**Proposed methods:**
Correlation analysis, pathway data extraction and association analyses. Depending on context information from individual CpG sites may be combined before or after association testing.

**Title: Fussy eating in infancy/toddlerhood and later child growth.**

**Supervisors/Collaborators:**
**Kate Northstone and Laura Howe**

**Background**:
Infant feeding can be a source of considerable stress for parents. Children who refuse to eat the 'right' foods or who appear extremely fussy can give particular cause for concern. Parents often worry that the child may be receiving insufficient nutrition. Parents of a fussy eater are likely to be concerned about the growth of their child, potentially supplementing their diet with energy-dense foods because of concerns that the child's diet may not meet their energy needs for growth and development. Rather than hamper growth, such actions could potentially result in fussy eaters being more likely to develop obesity. Only a handful of studies have been performed to date examining the associations between fussy eating and growth. These primarily cross-sectional studies provide conflicting evidence, have not looked at height growth, suffer from small sample sizes, and only one has been performed in the UK. Health visitors and other early-years practitioners are often the first port of call for parents worrying about their child's fussy eating, but they have little solid evidence-based advice to pass on to these worried parents

**Objective**:
The aim of this study is to use data collected by a large UK-based prospective cohort study (the Avon Longitudinal Study of Parents and Children (ALSPAC)) to assess the association between fussy eating (defined by parent-completed questionnaires at 15 months and 24 months) and height and BMI growth up to age 17 years.

**Methods**:
1. Carry out literature review on fussy eating in early childhood and growth
2. Prepare dataset, including the selection of appropriate confounding variables
3. Using data from the ALSPAC cohort, examine the association between fussy eating at ages 15 and 24 months and height, weight, and body mass index (BMI) up to and including 17 years of age. This analysis will use multilevel models, which have been previously developed by the supervisors
4. Write paper for submission to peer-reviewed journal

**Learning objectives:**
This project offers the student the opportunity to become familiar with the ALSPAC cohort, and to carry out research in the areas of nutrition and obesity. The student will also gain experience in the analysis of repeated measures data, with full guidance and support from supervisors. The student will carry out the analysis and write the paper up for publication in a peer-reviewed journal, and will therefore gain experience of the publication process.

**21**

**Title: Identifying genetic variants that influence facial morphology**

**Supervisors:**
**Lavinia Paternoster**

**Outline:**
Craniofacial morphology is highly heritable, but little is known about which genetic variants influence normal facial variation in the general population. In the Avon Longitudinal Study of Parents and Children (ALSPAC) participants have had their faces scanned, generating 3D high-resolution images. Landmarks have been identified on these images and 3D distances, angles, ratios, as well as principal components (which capture the majority of the variation) have been generated. A previous genome-wide association study in ALSPAC analysed 54 distances and 14 principal components and identified a genetic variant in the gene *PAX3* that influences the position of the nasion (bridge of the nose) (Paternoster et al., 2012). Since this paper, a new, larger genome-wide imputation (using the 1000 genomes reference) is available for the ALSPAC participants. In addition, new facial parameters have been produced. This project will analyse this new data in an attempt to identify additional genetic variants for facial morphology.
As there are so many potential facial parameters and that to analyse them all introduces a multiple-testing problem, first an analysis will be undertaken to estimate which parameters have the highest proportion of variation explained by genome-wide SNPs. This will be done using a technique called Genome-wide Complex Trait Analysis (GCTA) (Yang et al, 2010) and will identify which parameters show the best potential for the next step. The most promising parameters will then be taken forward to genome-wide association analysis. For any associations identified, replication will be sought from cohorts with similar data available, such as Visigen.

**Objectives:**
1. Investigate the distributions and correlations of the facial parameters that have been generated.
2. Carry out Genome-wide Complex Trait Analysis (GCTA) to estimate the proportion of phenotypic variance explained by genome-wide SNPs for each of the parameters.
3. Identify genetic variants for the most promising parameters using genome-wide association analysis.
4. Attempt to replicate findings in an additional cohort.
5. Write up results for publication.

**References:**
Paternoster L et al (2012) AJHG 90:478-85; Yang J et al (2010) Nat Genet. 42: 565-9.

**22**

**Title: Investigating epigenetic mediators of atopic dermatitis risk factors**

**Supervisors:**
**Lavinia Paternoster and Caroline Relton**

**Background:** Atopic dermatitis (AD), or eczema, is a common chronic skin disease with prevalence rates in developed countries of 15-30% in children and 2-10% in adults.[1] It commonly starts during infancy and frequently precedes or co-occurs with food allergy, asthma and rhinitis.
Early life exposures (pre- and post-natal) are thought to be important risk factors for AD. Intrauterine exposures, caesarean delivery, breast-feeding practices, timing of introduction of particular foods, antibiotic use and allergen exposures are all thought to be important.[2-4] There is some evidence that rather than complete avoidance of particular allergens, exposure to non-pathogenic microbes at key points in immune development, to build tolerance, may be the key to preventing AD.[5] Previous studies have shown maternal AD is a stronger risk factor for AD in the child than paternal AD.[6,7] One possible explanation of this is if there are important intrauterine exposures that influence risk of AD.
Epigenetic alterations (such as DNA methylation) are thought to be an important mechanism for mediating the effects of early life exposures on long term health. Therefore, this project will involve investigating whether methylation changes mediate the link between specific prenatal exposures and the subsequent development of AD in children using data from the Avon Longitudinal Study of Parents and Children (ALSPAC).

**Objectives:**

1. To carry out a literature review of early-life exposures for AD.
2. Carry out observational association analysis for these in the ALSPAC cohort.
3. To investigate whether any of the associations between early-life exposures and AD are mediated by methylation.
4. Attempt to replicate findings in additional cohorts.
5. Write up results for publication.

**Methods:** This project will use data from the Avon Longitudinal Study of Parents and Children (ALSPAC). This population cohort has extensive data on AD diagnoses and symptoms at several time-points throughout childhood. In addition to the phenotype data, there is also extensive information from questionnaires and other measures on the exposures to various risk factors.
As part of the ARIES (Accessible Resource for Integrated Epigenomic Studies) project, 1000 mother-child pairs from ALSPAC have genome-wide methylation data (from the HumanMethylation450 (HM450K) BeadChip) at several time points (for the children these are: at birth, age 7 and age 15-17). In the first instance DNA methylation in cord blood will be analysed in relation to a range of prenatal exposures. Any loci shown to differ in DNA methylation according to these AD-relevant exposures will be assessed in relation to subsequent AD in childhood. The loci could also be interrogated using DNA methylation data from later time points to assess whether the changes in methylation persist across childhood. Replication will be sought with collaborating birth cohorts with complementary data available, such as Generation R.

**References:** 1. Bieber T. NEJM 358:1483-94(2008); 2. Roduit C et al. Immunol 130:130-6 (2012); 3. Torley D et al. Clin Exp Dermatol 38:449-56(2013); 4. Flohr C & Yeo L. Curr Probl Dermatol 41:1-34(2011); 5. Williams HC. Vet Dermatol 24:3-9(2013);6. Esparza-Gordillo J et al. Curr Opin Allergy Clin Immunol 10:418-26(2010); 7. Moffatt MF & Cookson WO. Clin Exp Allergy 28,56-61(1998)

**Title: Does DNA methylation influence bone mineral density in children?**

**Supervisors:**
**Dr Celia Gregson, Prof Jon Tobias, Prof Caroline Relton**

**Background:**

Osteoporosis is common, often asymptomatic and therefore insidious. The consequences, primarily fractures, represent a significant societal burden, both through direct medical costs (now estimated to be more than £2 billion per annum in the UK), and important social sequelae (1). Approximately 1 in 2 women and 1 in 5 men aged over 50 years will fracture their hip, spine or forearm during their lifetime (2). Bone mineral density (BMD), a principle clinical predictor of fracture, increases during childhood growth and puberty until peak bone mass is achieved – this 'set point' determines the skeletal reserve to withstand osteoporosis inducing factors which occur over the remainder of adult life. Hence those factors determining peak bone mass accrual have consequences over the life course.

Genome-wide association studies (GWAS) have to date identified and replicated 59 bone metabolism genes thought to regulate BMD and fracture risk (3-5). However, these explain <5% of the phenotypic variance in BMD. Currently the contribution of epigenetic regulation to BMD is unknown. Epigenetic DNA modifications, such as methylation, are expected to regulate BMD development during growth and puberty and therefore ultimately influence peak bone mass accrual.

**Objectives:**

The study will be based upon data already collected as part of the Avon Longitudinal Study of Parents and Children (ALSPAC) and analyses will examine:

1. The relationship between genome-wide DNA methylation and BMD (measured by DXA) in 1000 children (male and female) at age 7-9 and at then again at age 15-17
2. How changes in the genome-wide DNA methylation profile during puberty relate to BMD
3. Whether methylation status of established BMD genes influences their effect on BMD before and during pubertal growth

**Methods:**

Genome-wide DNA methylation has already been analysed using the 450K Illumina methylation Platform as part of the Accessible Resource for Integrated Epigenomics Studies (ARIES) project, which is a sub-study of ALSPAC.

1. Review the literature on DNA methylation and bone growth in children
2. Define appropriate exposure and BMD outcome measures and agree an analysis plan with supervisors
3. Obtain relevant dataset from ALSPAC & supervisors
4. Complete cross-sectional statistical analyses
5. Write up for publication and presentation at inter/national conference

There is also scope to extend analyses to include (a) bone micro-architecture and geometry as measured by pQCT (peripheral quantitative computer tomography) and (b) the application of Mendelian randomization using established BMD SNPs to assess the direction of the causal relationship between DNA methylation and BMD.

**References**


1. Burge Rt WD. The cost of osteoporotic fractures in the UK: projections for 2000-2020. J Med Econ. 2001;4:51-62.
2. van Staa TP, Dennison EM, Leufkens HG, Cooper C. Epidemiology of fractures in England and Wales. Bone. 2001;29(6):517-22.
3. Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. Nat Genet. 2012;44(5):491-501.
4. Duncan EL, Danoy P, Kemp JP, Leo PJ, McCloskey E, Nicholson GC, et al. Genome-Wide Association Study Using Extreme Truncate Selection Identifies Novel Genes Affecting Bone Mineral Density and Fracture Risk. PLoS Genet. 2011;7(4):e1001372.
5. Paternoster L, Ohlsson C, Sayers A, Vandenput L, Lorentzon M, Evans DM, et al. OPG and RANK polymorphisms are both associated with cortical bone mineral density: findings from a metaanalysis of the Avon longitudinal study of parents and children and gothenburg osteoporosis and obesity determinants cohorts. Journal of Clinical Endocrinology Metabolism. 2010;95(8):3940-8.

**Title: Using metabolomic profiles to investigate the association between cardiovascular disease and osteoporosis**

**Supervisors: Dr Celia Gregson, Prof Jon Tobias, Prof David Evans, Prof Debbie Lawlor**

**Background:**

Cardiovascular disease (CVD) and osteoporosis are both common age-related conditions, usually progressing silently until serious clinical end-points such as myocardial infarction or fracture occur. Atherosclerosis, which underlies CVD, begins in young adulthood, with smooth muscle cells proliferation leading to intimal thickening (1, 2). Strong epidemiological evidence exists linking CVD with osteoporosis (and bone turnover), with both acting as exposure and outcome (3). Both conditions share common etiological factors, mediated through a range of differing metabolic pathways. Furthermore, individual pharmacological treatments may also be important, for example anti-resorptive therapy, which reduces bone turnover, is the principle and widely used osteoporosis treatment, and in animal models reduces atherosclerosis.

If CVD increases osteoporosis risk by increasing bone turnover, this will have important clinical implications, as CVD should then considered as an additional independent risk factor in patients being considered for anti-resorptive therapy for osteoporosis. Currently the role bone turnover plays in explaining the relationships between CVD and osteoporosis is being examined in peri-menopausal women assessed as part of the Avon Longitudinal Study of Parents and Children (ALSPAC). Newly available nuclear magnetic resonance assays now allow us to measure a wide range of metabolic phenotypes (4). Understanding the metabolomic profile common to CVD and osteoporosis, and the extent to which bone turnover contributes to this, is a novel approach to understanding cardiovascular-bone relationships.

**Objectives:**

This study will be based upon data already collected as part of the ALSPAC and fit within an ongoing investigation of the role bone turnover plays in explaining the link between CVD and osteoporosis. These analyses will specifically aim to:

1. identify the metabolomic profile common to both CVD (quantified by carotid artery intimal thickness) and osteoporosis (quantified by DXA BMD)
2. identify the metabolomic profile associated with bone turnover as measured by serum CTX
3. determine the extent to which the CTX-specific metabolomic profile explains the relationship between carotid intimal thickness and BMD

**Methods:**

6. Review the literature on nuclear magnetic resonance metabolomic profiling and the relationship between CVD and osteoporosis
7. Define appropriate exposure and outcome measures and agree an analysis plan with supervisors
8. Obtain relevant dataset from ALSPAC & supervisors
9. Complete statistical analyses
10. Write up for publication and presentation at inter/national conference

**References:**


1. Sinzinger H, Silberbauer K, Auerswald W. Does prostacyclin (PGI2) regulate human arterial intima smooth muscle cell proliferation in early atherogenesis? Blood vessels. 1980;17(1):58-60. PubMed PMID: 6986925.
2. Millonig G, Malcom GT, Wick G. Early inflammatory-immunological lesions in juvenile atherosclerosis from the Pathobiological Determinants of Atherosclerosis in Youth (PDAY)-study. Atherosclerosis. 2002 Feb;160(2):441-8. PubMed PMID: 11849669.
3. den Uyl D, Nurmohamed MT, van Tuyl LHD, Raterman HG, Lems WF. (Sub) clinical cardiovascular disease is associated with increased bone loss and fracture risk; a systematic review of the association between cardiovascular disease and osteoporosis. Arthritis Res Ther. 2011;13(1). PubMed PMID: WOS:000289523500016. English.
4. Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lyytikainen LP, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet. 2012 Mar;44(3):269-76. PubMed PMID: 22286219. Pubmed Central PMCID: 3605033.

**Title: Metabolomic profiles in individuals with an extreme High Bone Mass phenotype**

**Supervisors: Celia Gregson, Jon Tobias and David Evans**

**Background:**
We have previously collected a unique and extreme High Bone Mass (HBM) population; having screened 335,115 historical DXA (bone density) scans across 13 UK NHS centres to identify individuals with a bone mineral density (BMD) ≥+4 SDs above normal. HBM, with prevalence of 0.18%, was found to be associated with clinical characteristics suggestive of a mild skeletal dysplasia (e.g. mandible enlargement, extra localized bone) as well as increases in adiposity (1, 2). Whilst a handful of HBM cases have subsequently been identified as having activating *LRP5* mutations causing HBM, the genetic etiology is current yet to be established in the vast majority (and forms part of an ongoing project).

Newly available nuclear magnetic resonance assays now allow us to measure a wide range of metabolic phenotypes (3). Metabolomics offers a novel methodology with which to characterize HBM, and identify physiological processes which may elucidate the underlying etiology of extremely high BMD.

Declining bone mineral density (BMD) results in osteoporosis. BMD is a principle clinical predictor of fracture. Osteoporosis is common and the consequent fractures represent a significant societal burden, through both direct medical costs (now estimated >£2 billion per annum in the UK), and important social sequelae (4). Understanding the metabolic regulation of extreme BMD is intended to have wider implications for the general population.

**Objectives:** This study will be based upon data already collected as part of the High Bone Mass Study. These analyses will specifically aim to define the metabolomic profile in a rare population (n=355) with an extreme high bone mass phenotype and compare this with two control populations; firstly first-degree relatives who have normal BMD, and secondly a general population control group, selected from the Avon Longitudinal Study of Parents and Children (ALSPAC), in which case analyses will be restricted to females only.

**Methods:**

1. Review the literature on nuclear magnetic resonance metabolomic profiling and High Bone Mass
2. Define appropriate exposure and outcome measures and agree an analysis plan with supervisors
3. Obtain relevant dataset from ALSPAC and supervisors
4. Complete statistical analyses
5. Write up for publication and presentation at inter/national conference

**References**

1. Gregson CL, Steel SA, O'Rourke KP, Allan K, Ayuk J, Bhalla A, et al. 'Sink or swim': an evaluation of the clinical characteristics of individuals with high bone mass. Osteoporos International. 2012;23(2):643-54.

2. Gregson CL, Paggiosi MA, Crabtree N, Steel SA, McCloskey E, Duncan EL, et al. Analysis of body composition in individuals with high bone mass reveals a marked increase in fat mass in women but not men. The Journal of clinical endocrinology and metabolism. 2013;98(2):818-28.
3. Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lyytikainen LP, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet. 2012 Mar;44(3):269-76. PubMed PMID: 22286219. Pubmed Central PMCID: 3605033.
4. Burge Rt WD. The cost of osteoporotic fractures in the UK: projections for 2000-2020. J Med Econ. 2001;4:51-62.

**Title: Genetic influences on the brain reward system and teenage alcohol drinking**

**Supervisors: Dr Luisa Zuccolo and Dr Jon Heron**

**Background:**
To date, very few common genetic variants are known to be associated with alcohol behaviour in populations of European ancestry (1). In particular, the genetic component of dimensions of teenage drinking has not been explored yet, despite suggestions that alcohol behaviour has a strong heritable component even at a young age.

Based on a recent functional characterisation study, a common haplotype in the RASGRF2 gene is thought to affect alcohol drinking through effects on the brain reward system, i.e. by influencing dopamine release following ethanol ingestion, but to have no effect on ethanol metabolism or sensitivity (2). This 'propensity to addiction' could explain an association with number of drinking episodes observed in a small sample of 700 sixteen year old subjects (2). This genetic variant is a very promising candidate for explaining some of the variability in alcohol behaviour at the population level. Replication of the initial findings in a large population-based cohort is required to obtain robust evidence of its involvement in the susceptibility to problem drinking.

**Objective:**
To examine the association of the RASGRF2 haplotype with teenage alcohol drinking in participants from the Avon Longitudinal Study of Parents and Children (ALSPAC).

Detailed information on alcohol consumption throughout the teenage years is available, from which several alcohol phenotypes and their trajectories over time have been derived.

**Methods:**

1. Review the literature on RASGRF2 genetic variation and alcohol behaviour
2. Identify relevant alcohol phenotypes and agree an analysis plan with supervisors
3. Obtain the relevant data from ALSPAC databases/supervisors
4. Complete statistical analysis
5. Write up for publication

**References:**

(1) Schumann G, et al. (2011) Genome-wide association and genetic functional studies
identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. Proc Natl Acad Sci USA 108(17):7119–7124.

(2) Stacey D et al. RASGRF2 regulates alcohol-induced reinforcement by influencing mesolimbic dopamine neuron activity and dopamine release. Proc Natl Acad Sci U S A. 2012 Dec 18;109(51):21128-33.

**Title: A Mendelian randomization study to identify nutrients which cause prostate cancer incidence and progression**

**Supervisors: Sarah Lewis, Carolina Bonilla, Richard Martin**

**Background:**
Based on our previous work and other research which has been carried out in this area we hypothesis that milk, calcium, vitamin D, folate and vitamin B12 increase the risk of prostate cancer or prostate cancer progression, and that selenium, vitamin E and lycopene are protective. The objective of this mini-project is to investigate whether one of these nutrients is causally related to prostate cancer using a Mendelian randomization design, in a large consortium of prostate cancer case-control studies or cohort studies.
The plan is to make use of single nucleotide polymorphism data which is available within the PRACTICAL consortium, a consortium of 21 epidemiological studies comprising of over 22,000 cases and an equivalent number of controls, to analyse SNPs related to nutrient intake, synthesis, metabolism, transport or cellular uptake for associations with prostate cancer risk. It will then be possible make inferences about the related exposure on prostate cancer risk and progression based on the associations (or lack of) between the SNPs and outcome. This will be the largest study of its type and will give a definitive insight into the nutrients casually related to prostate cancer, which may lead to public health interventions which aim to modify diet and reduce prostate cancer risk.

The project will involve training in the following areas:
Basic genetics and epidemiology analysis, meta-analysis, mendelian randomization.

**Title: Foetal Testosterone, 2D:4D Digit Ratio, and Social Cognition**

**Supervisors: Marcus Munafo**

**Background:**

Foetal testosterone, assessed indirectly through 2D:4D digit ratio, is reportedly associated with numerous mental health outcomes, including, depression, psychopathy, anxiety and ADHD [1-4]. One group of outcomes that have attracted considerable research interest are related to empathy and autism [5]. However, findings in this area are contentious – many studies have employed small sample sizes, and the validity of 2D:4D as a simple retrospective biomarker of testosterone exposure has been questioned [6].

ALSPAC has measures of 2D:4D from over 7,000 children at 11yrs, and considerable information about social communication and emotional processing ability during adolescence (e.g., SCDC, DANVA, emotional triangles). Genetic variants associated with 2D4D have been identified, which provide an instrument for Mendelian randomisation analyses. This dataset therefore provides the opportunity for a definitive test of associations between 2D:4D and emotional processing.

This project will therefore investigate relationships between 2D:4D and emotional processing, both observationally and, causally, using Mendelian Randomisation. We will analyse links associations between phenotypic information for emotional processing ability and 2D:4D and, using MR, test for causal relationships between 2D:4D and emotional processing during adolescence.

1) Austin EJ, et al. *Personality and Individual Differences* 2002; 33: 1115-24.
2) Bailey A, Hurd P. *Personality and Individual Differences* 2005; 39: 829-36.
3) Blanchard A, Lyons M. *British Journal of Forensic Practice* 2010; 12: 23.
4) McFadden D et al. *Clinical Neuroscience Research* 2005; 5: 233-45.
5) Manning JT et al. *Medicine and Child Neurology* 2001; 43: 160-4.
6) Medland SE et al. *American Journal of Human Genetics* 2010; 86: 519-25.

**Title: Effects of 5-HTTLPR and Early Childhood Adversity on Cognitive and Emotional Processing**

**Supervisors: Marcus Munafo**

**Background:**

There is ongoing debate as to whether the risk associated with psychosocial adversity on depression is moderated by specific genetic polymorphisms, particularly the 5-HTTLPR. It has been suggested that deficits in cognitive and emotional processing may act as transdiagnostic intermediate markers for anxiety and depression.

In a small sample (n = 238) of adolescents, Owens and colleagues [1] recently reported evidence that genetic polymorphism in the 5-HTTLPR, and exposure to childhood adversity, moderated emotional processing in adolescence. Emotional processing was assessed using the Affective Go No Go and Probability Reversal Learning tasks. They concluded that these emotional processing differences may serve as endophenotypes for anxiety and depression in those individuals with genetic susceptibility who were also exposed to childhood adversity.

Small candidate gene studies, and studies testing for gene by environment interactions, have been found to be unreliable, producing evidence of implausibly large genetic effects [2]. Given the small size of their sample, and the controversy surrounding the evidence for role of an interaction of 5-HTTLPR with adversity in depression, it seems important to test the association reported by Owens and colleagues in a larger, more powerful study design.

We propose to test the hypothesis that the risk associated with psychosocial adversity on depressive endophenotypes is moderated by specific genetic polymorphisms in ALPSAC. We will test this hypothesis using the 5-HTTLPR variant data, measures of childhood adversity, and data from the Affective Go No Go and Probability Reversal Learning tasks which were completed at age 18 years.

1) Owens M, et al. *PLOS One* 2012; 7: e48482.
2) Flint J, Munafò MR. *Current Opinion in Neurobiology* 2013; 23: 57-61.

**30**

**Title: Inhibitory Control Training and Alcohol Consumption**

**Supervisors: Marcus Munafo**

**Background:**

The detrimental effects of alcohol on inhibitory control are well reported [1], and are believed to underlie many of the risky behaviours associated with acute alcohol consumption including aggression and unsafe sex. Recent research suggests that inhibitory control can be improved using computer-based training paradigms [1-4]. These findings offer a potential avenue for intervention in alcohol drinkers who find it difficult to control drinking and who engage in risky behaviours when intoxicated [5]. However, it remains unclear whether training effects generalise to other (non-trained) aspects of inhibitory control, other stimuli and other relevant behaviours.

The current project will investigate the effects of single session inhibitory control training in heavy social alcohol consumers on general and alcohol-specific inhibitory control tasks. We will use a standard inhibitory control training (ICT) training task that is based on a go no-go paradigm and which has been previously shown to be effective in smokers (unpublished data). We hypothesise that individuals randomised to receive active (compared to placebo) training will show reductions in post-training inhibitory control to both neutral and alcohol-related stimuli. In order to assess whether effects generalise, we will include another test of inhibitory control (stop-signal task) and a taste test to measure post-training consumption of alcohol. In addition, participants will be required to complete an alcohol consumption diary for one week post-training.

This pilot study will examine the potential efficacy of ICT to improve inhibitory control in heavy drinkers. It will inform the design of future studies and establish the extent to which training generalises to other stimuli (alcohol vs. non-alcohol), and to other relevant behaviours (e.g., actual alcohol consumption). High levels of acute alcohol consumption are common in the UK and result in a number of personal and social harms. While alcohol control strategies aim to reduce drinking at a population level, many young drinkers still regularly drink to excess and ICT may offer a novel approach to reduce alcohol-related harms in non-abstaining drinkers.


1) Field M, et al. *Alcoholism: Clinical and Experimental Research* 2010; 34: 1346-52.
2) Houben K, et al. *Appetite* 2011; 56: 345-9.
3) Berkman ET, et al. *Journal of Neuroscience* 2014; 34: 149-57.
4) Motes MA, et al. *Brain and Cognition* 2013; 25: 44-62.
5) Bowley C, et al. *International Journal of Psychophysiology*; 89: 342-8.

**31**

**Title: Does BMI influence Smoking Behaviour? A Mendelian randomisation study**

**Supervisors: Marcus Munafo**

**Background:**

There is good evidence that smoking lowers body weight, and that this may be due to the effects of nicotine on appetite and energy expenditure [1]. However, it is likely that the relationship between smoking and body mass index (BMI) is more complex. For example, some smokers report smoking in order to aid weight loss [2], so it is possible that the association is bidirectional, that is to say that as well as smoking influencing BMI, BMI influences smoking behaviour. As smoking and BMI are both strongly socially patterned and are associated with a range of other lifestyle factors, the relationship between them is difficult to evaluate using conventional epidemiological methods.

This problem of confounding can be minimised using Mendelian randomisation analyses, whereby genetic variants associated with BMI are used as proxies for measured BMI. Recently, a large study has demonstrated associations between BMI variants and smoking initiation and heaviness, suggesting either that BMI influences smoking or perhaps that there is a shared biological basis for both tobacco and food intake [3]. However, the smoking phenotypes used in this study were limited e.g. smoking initiation was based on reports of never and ever smoking.

This project will investigate the association between BMI and a range of smoking behaviours (initiation/heaviness/cessation) in the ALSPAC cohort (in both mothers and children) by running both observational analyses and Mendelian randomisation analyses. ALSPAC has detailed measures of smoking behaviour, including cotinine and trajectories of smoking initiation [4]. For the Mendelian randomisation analysis, an allelic risk score for BMI can be constructed from 32 single nucleotide polymorphisms (SNPs) [3].

1) Freathy R, et al. *International Journal of Epidemiology* 2011; 40: 1617-28.
2) Chiolero A, et al. *American Journal of Clinical Nutrition* 2008; 87: 801-9.
3) Thorgeirsson T, et al. *Translational Psychiatry* 2013; 3: e308.
4) Heron J, et al. *Nicotine and Tobacco Research* 2011; 13: 1266-75.

**Title: Acute Anxiety and Multiple Object Tracking**

**Supervisors: Marcus Munafo**

**Background:**

Anxiety disorders are associated with a number of cognitive changes, including hypervigilance to threat, which can have detrimental effects on other cognitive processes such as working memory [1]. These effects, however, are not restricted to anxiety disorders, but may also occur when anxiety levels are raised in non-disordered individuals [1]. Understanding these effects is important as we are often required to perform in pressurised situations in which anxiety is raised, as is the case in many professional occupations including medicine, sport and aviation.

Much of our understanding of the effects of anxiety on human behaviour and cognition comes from research using patient samples or individuals high in trait anxiety, and which often measure disorder relevant outcomes (e.g., face processing in social anxiety disorder). Laboratory research is limited due to a lack of reliable models of anxiety induction. The Tobacco and Alcohol Research Group has a programme of research investigating the effect of acute anxiety on human cognition and behaviour using the 7.5% carbon dioxide ($CO_2$) inhalation procedure, which temporarily increases physiological and subjective indices of anxiety in most people. Using this model, it has been shown that acute anxiety alters threat processing [2, 3], face perception [4] and speech perception [5].

The proposed study will test the effects of $CO_2$-induced anxiety on attention allocation using a validated multiple object tracking task. The primary objective is to investigate whether anxiety affects attention by narrowing focus on fewer objects, or broadens focus so that distraction is more likely. We will use a continuous measure of accuracy (degree of error in estimated object trajectory) alongside a measure of hand coordination. According to the narrowing hypothesis, we should see a greater impact of anxiety on less attention demanding objects compared with those more attention demanding objects. In contrast, the broadening hypothesis predicts that all objects will suffer when the participant is anxious. We will recruit 30 healthy volunteers (50% male) to complete one session of two 20-minute inhalations (7.5% $CO_2$ and normal air).

The findings from this study will further our understanding of attention allocation during acute anxiety. An attentional focus on threat is well reported during acute anxiety [1, 2, 3] and in anxiety disorders [6], but this study will utilise non-threatening stimuli and thereby test attention allocation more generally. We will test predictions from two conflicting hypotheses that will enable us to build a model of attention allocation during acute anxiety. These findings will have implication for real world situations that require specific allocation of attention under stress.

6) Robinson OJ, et al. *Frontiers in Human Neuroscience* 2013; 7: 203.
7) Garner M, et al. *Neuropsychopharmacology* 2011; 36: 1557-62.
8) Cooper R, et al. *Cognition and Emotion* 2012; 27: 273-82.
9) Attwood AS, et al. *Psychological Science* 2013; 24:1591-4.
10) Mattys SL, et al. *Psychological Science* 2013; 24: 1606-8.
11) Cisler JM, et al. *Clinical Psychology Review* 2010; 30: 203-16.

**Title: Development of an Automated Drinking Rate Monitor**

**Supervisors: Marcus Munafo**

**Background:**

Studies of drinking behaviour often require accurate measurement of the rate at which people consume a drink [1-2]. Currently, these measurements are obtained by recording video footage of a person drinking, and then manually coding the number and size of sips taken in a given time frame [1-2]. This method is extremely labour intensive, and can introduce noise to measurements through coding errors and inaccuracies. A considerable improvement would be to use a dedicated drink rate monitoring device (unobtrusively concealed so as not to alter drinking behaviour) that automatically measures and records drinking rate, perhaps by detecting when the glass is lifted and changes in the weight of liquid in the glass. The aim of this project is to design, build and validate such a device.

Recent advances in commercially available, ultra compact, low-power, generic technologies (e.g., TinyDuino, TinyLily) bring new opportunities for developers of this type of device. These systems are modular in nature, allowing the size and power conscious designer to select only the components essential for an optimal solution. For example, a designer may select a sensor component that can detect when the glass is lifted and the weight of liquid in a glass, a processor component that can translate these into drinking measures, and a networking component that can transmit this information to an app running on a smartphone or laptop nearby, while not selecting other options that consume space and power (e.g., USB connectivity). Together, the selected components would be small enough to fit within a space a few centimetres in diameter by a few millimeters high, and could be powered by a single coin battery, which means they could potentially be integrated into a modified drinking glass or drink mat.

This project, which will be supported by both the Faculty of Engineering, and the School of Experimental Psychology, will involve the design and building of an automated drink rate monitor device, the design and writing of application software necessary to make a range of drink rate measures available to researchers (e.g., volume consumed per time period, number of drinks per time period), and the validation of the monitor device and application in an experimental setting. Research into drinking behaviour is likely to grow in response to the increasing concern about the personal and societal effects of high levels of alcohol consumption. A device of this kind will undoubtedly be of considerable utility in future studies of this kind.

1) Attwood A, et al. *PLOS ONE* 2013; 7: e43007.
2) Higgs S, et al. *Cognitive and Behavioural Effect* 2008; 43: 630-35.

**34**

**Title: Early Life Stress, Age of Puberty and Reproductive Behaviour**

**Supervisors: Marcus Munafo**

**Background:**

Early life experience (ELE) has profound effects on later outomes. For example, links between adverse childhood environments and early sexual maturity in girls are reasonably well-established, while aggressive behaviour in adolescence and beyond is associated with both physical and social deprivation in childhood [1-3]. A life history perspective on these findings proposes that the allocation of energetic resources to reproductively important factors (e.g., sexual maturity, growth trajectories) is determined by flexible mechanisms that function to increase reproductive success, and that these outcomes may be functional responses to the environment, rather than reflecting pathology [4].

Early life adversity (characterized by unpredictable or harsh environments) may therefore predispose individuals to a 'fast' life history strategy – in hostile environments with high extrinsic mortality and morbidity, it pays to prioritise reproduction early in life, and to trade offspring quality for number. These strategies are characterised by early sexual maturity and reproduction in girls, and high levels of risk taking (sexual and otherwise) and aggression in both sexes. Earlier work in this area, although extensive, is often fragmented, looking at limited individual outcomes as a response to particular adverse events (e.g., early menarche in response to father absence). Where more multidimensional data have been employed, there has been a tendency to produce composite measures of environmental harshness. These techniques fail to determine why certain exposures lead to specific later outcomes (causal MR flannel here).

This study will provide a multi-faceted understanding of the causal variables underlying behaviour and the psychological correlates that form and mediate individual life history strategies, allowing a strong test of hypotheses of psychosocial acceleration in response to early adversity. We will use latent variable analysis and structural equation modelling approaches to investigate the relationships between ELE and both physiological (e.g., age at menarche/puberty, growth trajectories, size) and psychological/behavioural (e.g., sexual/aggressive behaviours; emotional well-being) outcomes in adolescence. Outcome data in adolescence will be used to measure both physical and psychological health, as well as providing proxy measures of reproductive strategy (e.g., pubertal development, sexual activity) and wider social behaviour. Genetic variants associated with cortisol levels will be used as an instrumental variable for stress exposure in Mendelian randomisation analyses, to establish causality.

1) Belsky J, et al. *Child Development* 2007; 78: 1302-21.
2) Nettle D, et al. *Developmental Psychology* 2012; 48: 718-21.
3) Simpson JA, et al. *Developmental Psychology* 2012; 48: 674-86.
4) Chisholm, JS. *Current Anthropology* 1993; 34: 1-24.

**Title: Self-Titration of Nicotine in E-Cigarette Users**

**Supervisors: Marcus Munafo**

**Background:**

Use of electronic cigarettes ('e-cigarettes') has rapidly increased in recent years, and there are now an estimated 1.3 million current e-cigarette users in the UK [1]. Initial studies suggest that these devices may be effective smoking cessation aids [2]. However, their efficiency in terms of delivery of nicotine (the primary addictive constituent of tobacco) has been questioned [1,3,4].

It is well-established that tobacco smokers modify their smoking behaviour to self-titrate nicotine to a level appropriate to their need [5,6]. This compensatory behaviour is achieved, for example, through varying puff volume and number of puffs taken per cigarette. The goal of this study is to determine if the same compensatory behaviour is observed in e-cigarette users. Specifically, we will examine potential plasticity of e-cigarette use in response to nicotine content of e-cigarette solution.

E-cigarette vapour inhalation patterns will be examined in relation to e-cigarette nicotine solution content (0.6 mg/ml, 1.2 mg/ml, 1.8 mg/ml and 2.4 mg/ml) using a double-blind within-subjects design. An e-cigarette with in-built puff counter will be used to assess inhalation patterns over the course of 24 hours. All participants will be current, daily smokers recruited from the student body/general population. This study will inform a larger laboratory-based study, incorporating collection of biological samples.

1) Action on Smoking and Health (2013). http://ash.org.uk/files/documents/ASH_715.pdf
2) Bullen C, et al. *Lancet* 2013; *382:* 1629-37.
3) Vansickel AR, et al. *Cancer Epidemiology, Biomarkers and Prevention* 2010; 19: 1945-53.
4) Goniewicz M, et al. *Nicotine and Tobacco Research* 2013; 15: 158-66.
5) Strasser AA, et al. *Drug and Alcohol Dependence* 2007; 86: 294-300.
6) McNeill A, Munafo MR. *Journal of Psychopharmacology* 2013; 27: 13-8.

**36**

**Title: Preventing Health Warning Avoidance in Regular Smokers**

**Supervisors: Marcus Munafo**

**Background:**

Health warnings on cigarette packs provide important information to smokers about the health consequences of smoking, and research shows that attention to health warnings leads to meaningful changes in behaviour, such as making quit attempts [1]. Previously, using eye-tracking technology, we have found that plain packaging of cigarettes increases visual attention to health warnings among adult and adolescent non-daily smokers but not among daily smokers [2-3]. More recently, we have found that daily smokers actively avoid cigarette pack health warnings [4], presumably through top-down cognitive control. Other studies have also shown health warning avoidance among smokers [5] using self-report measures. However, our study is the first to assess avoidance with objective eye-tracking measures. Research should now focus on increasing attention to health warnings among smokers.

Although cigarette packs in the United Kingdom display the health warning on the lower half of the pack, many countries require warnings to be placed on the top half of the pack, as this is the point where people typically start reading. It may also be the case that daily smokers in the UK have learned to look away from the lower half of the pack, where they know the health warning is located. For these reasons, placing health warnings on the upper half of cigarette packs might be an effective method of increasing attention to warnings. The aim of this project is to determine whether moving the location of the health warning on cigarette packs leads to an increase in attention to the warning.

We will test this using eye-tracking technology. Non-smokers and daily smokers will be shown a series of branded cigarette packs, with health warnings located either on the lower or upper half of the pack. We will record the number of eye movements participants make to the health warning and branding on the two pack types. The student will be taught how to use the eye-tracker, and how to run experiments in MATLAB. Empirical data on which position elicits greatest attention to the health warnings will be important in developing and refining tobacco control policies.

1) Hammond D, et al. *American Journal of Preventive Medicine* 2007; 3: 202-9.
2) Munafo MR, et al. *Addiction* 2011; 106: 1505-10.
3) Maynard OM, et al. *Addiction* 2013; 108: 413-9.
4) Maynard OM, et al. *Drug and Alcohol Dependence* 2014.
5) Hammond D, et al. *American Journal of Public Health* 2004; 94: 1442.

**Title: Social Learning Mechanisms in Social Anxiety**

**Supervisors: Marcus Munafo**

**Background:**

Social anxiety (SA) is common, debilitating and contributes significantly to the global burden of disease [1]. Nearly half of all patients fail to respond to current treatments [2]. A better understanding of the mechanisms which maintain SA is required to develop new interventions, and identify which patients are likely to respond to specific treatments.

We have developed a probabilistic Social Learning Task which assesses the ability to learn social rules such as 'this person likes me' and 'that person dislikes her' based on online learning of response-feedback contingencies. We found that socially anxious individuals automatically favoured learning negative self-referential cues making them hypersensitive to learning when they were disliked and identifying social learning as a promising target for intervention [3].

We are currently running an fMRI study is to describe the brain mechanisms involved in normal social learning and so identify the abnormal processes involved in SA. Activity in the anterior rostral medial frontal cortex (arMFC) is associated with self- and other-knowledge, and inferring mental states to others [4]. We hypothesise therefore that normal social learning will be associated with activity in this region. Hyperactivity in the amygdala is implicated in maintaining SA via selectively biasing attention towards threat [5]. We hypothesise that the social learning abnormalities in SA will be associated with increased amygdala activation and mediated via selective attention to cues signalling that the self is disliked.

The aim of this proposed study is test the latter part of this hypothesis, namely that learning about the self in social anxiety is associated with biased attention towards negative self-referential information. We propose to test this hypothesis in an analogue sample selected for high (n = 24) and low (n = 24) SA using the Social Learning Task in conjunction with simultaneous eye-tracking. Threat-selective attention will be assessed using eye-tracking to measure the direction of the first saccade and dwell time spent looking at positive/negative cues during the task.

The results of this project will inform the next steps in translation to clinical application. Paradigms which modify attention for therapeutic gain are well established [6], but their efficacy and mechanisms of action in SA are poorly understood. If, for example, we find that social learning in SA is mediated by attention then the next step is to develop an attention retraining intervention which targets social learning.

1) Wittchen HU, et al. *European Neuropsychopharmacol* 2011; 21: 655-79.
2) Acarturk C, et al. *Psychological Medicine* 2009; 39: 241-54.
3) Button KS, et al. *Journal of Behavior Therapy and Experimental Psychiatry* 2012; 43: 1082-7
5) Amodio DM, et al. *Nature Reviews Neuroscience* 2006; 7: 268-77.
6) Shechner T, et al. *Depression and Anxiety* 2012; 29: 282-94.
7) Bar-Haim Y. *Journal of Child Psychology and Psychiatry* 2010; 51: 859-70.

**Title: Investigating links between inflammatory markers and aggressive behaviour using a Mendelian randomisation approach**

**Supervisors: Marcus Munafo**

**Background:**

Evidence from both animal and human studies suggests that inflammatory markers are associated with aggressive behavioural traits. In a recent case control study, individuals with intermittent explosive disorder, a condition of impulsive aggression, demonstrated elevated levels of C-reactive protein (CRP) , an acute phase protein, and interleukin-6 (IL-6), an inflammatory cytokine, compared to healthy controls and controls with other psychiatric conditions [1]. However, it is unclear from these studies whether inflammatory markers are a cause or a consequence of aggressive behaviour, or whether this association is due to confounding by other lifestyle factors.

This project will investigate whether elevated inflammatory markers play a causal role in aggressive behaviour using Mendelian randomisation methods. Mendelian randomisation minimises the problems of confounding and reverse causality in observational studies by using genetic variants associated with exposures of interest as proxies for measured exposures [2]. Genetic variants have been identified which are robustly associated with CRP and IL6 levels; these have been used previously to investigate the potential role of these inflammatory markers in cardiovascular disease [3,4]. The associations of these variants with measures of aggressive behaviour will be explored in the ALSPAC children. Measures of aggressive behaviour and conduct disorder have been collected on a number of occasions in ALSPAC during childhood and adolescence [5,6].

1) Coccaro EF, et al. *JAMA Psychiatry* 2014; 71:158-65.
2) Davey Smith G, et al. *Int J Epidemiol.* 2003; 32: 1-22.
3) CRPCHDG Collaboration et al. *BMJ* 2011; 342: d548.
4) Interleukin-6 Receptor MR Consortium. *Lancet* 2012; 379: 1214-24.
5) Kretschmer, T, et al *Journal of Research on Adolescence* 2014; 24: 177-85.
6) Heron J, et al. *Addiction* 2013; 108: 2129-38.

**39.**

**Title: Investigation of cell mixture adjustments in analysis of DNA methylation in cord blood and children's blood**

**Supervisors: So-Youn Shin, E. Andres Houseman, Tom Gaunt, and Caroline Relton**

**Background:**

DNA methylation is a chemical modification to the cytosine residues in DNA that affects the gene expression. The change responds to environmental factors and is believed to be heritable. Recent technology enabled us to study genome-wide DNA methylation data in human measured on high throughput arrays, such as Illumina 450k microarray, in a relatively standardized manner. Such data opens a new possibility to identify epigenetic markers for diseases or disease-related traits, however, its analysis is not straightforward.

One of the challenges is due to cellular heterogeneity. Blood or any other tissue sources is a cell mixture where each cell has different methylation profile. It has been shown that methylation level differences at some CpG sites are due to the cell mixture variability[1]. Currently there are a couple of reference-based and reference-free methods adjusting for cell mixture[1,2,3,4]. However, all reference-based methods have been developed based on adults' blood. Given that the cell mixture changes across age[4], it is critical to investigate how these methods perform in cord blood and children's blood as well. Such methods require evaluation in the context of a specific question. This project will utilise data on cholesterol levels at 3 different ages (birth, in childhood and in adulthood) and evaluate the relationship between cholesterol levels and DNA methylation after applying the different cell mixture adjustment approaches.

ALSPAC has all the datasets available for this study, i.e. DNA methylation measured from cord bloods, children's bloods and adults' bloods, and cholesterol measures taken at the same time. Our study will provide a useful source of information to those who study the role of methylation at birth or in childhood.

Objectives:

1. Review the literature on 1) cell mixture of cord blood and children's blood in comparison to adults' blood and 2) current cell mixture adjustments in analysis of DNA methylation
2. Perform cross-sectional epigenome-wide association studies for cholesterol at birth, in childhood and in middle age under three different models: 1) no cell mixture adjustment, 2) reference-based adjustment, and 3) reference-free adjustment
3. Compare the results and the computational efficiency of the above three models
4. Write up for publication

References:

1) Houseman EA, et al. (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*

2) Houseman EA, et al. (2014) Reference-Free Cell Mixture Adjustments in analysis of DNA methylation data. *BMC Bioinformatics*
3) Zou J, et al. (2014) Epigenome-wide association studies without the need for cell-type composition. *Nature Methods*
4) Jaffe AE and Irizarry RA (2014) Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biology*