ISSN 2397-9917

Issue 7
Spring 2018

Galton Institute

Exploring Human Heredity

The Galton Review

CONTENTS

Front Cover Image: Rohan Yesudian, The King's School Chester

Published by:

The Galton Institute, 19 Northfields Prospect, London, SW18 1PE

Tel: 020 8874 7257 www.galtoninstitute.org.uk

General Secretary: Mrs Betty Nixon

executiveoffice@galtoninstitute.org.uk

Review Editor: Mr Robert Johnston

EDITORIAL

Following last year's exceptionally successful conference on *'Exploring Galton's Legacy'*, one might think it impossible for us to match it. However, this year's conference promises to be even more exciting as we explore the rapidly evolving subject of *'Genome Editing'.* The leading figure in what has come to be known as the CRISPR Revolution is, of course, Professor Jennifer Doudna, FRS of Berkeley California and we are thrilled that she has agreed to deliver this year's Galton Lecture. A huge amount of work has gone into the organisation of the day and the various speakers will consider the history of the subject along with its current and future applications. I hope you can join us at the Royal Society on Wednesday 31st October. Full details can be found on the website and instructions on how to obtain tickets are on page 16.

In this issue, we also have the winning essay from this year's schools' competition. The subject was "The role of statistics in medical and scientific research, especially in genetics" and the winner was Rohan Yesudian from the King's School in Chester. Many congratulations to a young man with a very bright future.

The Galton Institute is always pleased to sponsor various events around the world and this time we have reports from conferences in Cambridge and Brisbane which are probably about as far apart as two venues can be. Both reports make for interesting reading. If you are considering running a conference or workshop which is relevant to our aims then have a look at our website which provides details of how to apply for a grant towards the running of your event.

Finally, I have produced a book review of the latest biography of Charles Darwin, this time by the columnist and sometime biographer, A.E. Wilson. If you have also read it, please let me know if you agree with my findings or not.

Robert Johnston

2018 Essay Competition

This year's competition for 6th formers was based around statistics to mark the centenary of RA Fisher's iconic paper "The correlation between relatives on the supposition of Mendelian inheritance" in which he introduced the term 'variance'. Fisher is regarded by many as one of the greatest biologists who almost single-handedly created the foundations for modern statistical science.

Congratulations go to this year's winner, Rohan Yesudian from the King's School Chester whose essay is shown below.

****

"The role of statistics in medical and scientific research, especially in genetics"

Statistics is the tool to draw conclusions in scientific research under uncertainty. The phrase "statistically significant" is the Holy Grail for scientists – the acid test that determines whether their work is published or ignored. The power of statistics helps us validate findings; which in turn leads to advancements, such as new medical treatments. More recently, the use of complex statistical tests has led to the development of sophisticated algorithms through bioinformatics, which has helped us gain a better understanding of polygenic interactions within our genome.

Statistics has revolutionised the field of genetics. With over three billion base pairs of DNA nucleotides in the human genome, locating desirable genes for analysis is impractical. The use of association mapping has allowed us to locate these genes by identifying phenotypes, which give an indicator as to what genes are expressed. Linkage disequilibrium can help us understand the relationship between genetic markers and phenotypes. It is the difference between the observed frequency of alleles at particular loci on the genome and the expected frequency at these loci due to random evolutionary distribution which provides this information. If the frequency of the alleles at a particular locus is much different to the expected frequency, its arrangement is said to be independent of random evolutionary distribution and could have a particular phenotype - such as an hereditary disease. A simple statistical test for independence is the Pearson's Chi –squared test, which measures the 'fit' of the observed distribution of units with its

expected distribution. The value of χ^2 obtained from the test allows us to measure the deviation from the Hardy-Weinberg equilibrium (the frequency of alleles ignoring evolutionary variation – in effect, the null hypothesis) and allows us to conduct a hypothesis test to test whether it is statistically significant enough to have a causal effect on the phenotype. This issue of significance is particularly important as biological specimens show a lot of variation even with almost identical genetic makeup. This is why a confidence interval typically of 95% must be obtained through many repeated experiments to be sure there is a definitive relationship between the genotype and the phenotype.

If two alleles are arranged linearly along the same chromosome, they can be inherited as a haplotype (literally 'haploid genotype'). These can arise as short tandem repeat allele mutations (STRs) or a single nucleotide polymorphism mutation (SNP – referred to as a "haplogroup"). Determining haplotypes involves looking at multiple genomes, which may not be available. Therefore, a haplotype must be inferred using complex algorithms such as expectation-minimization, which models the missing data by simulating different designs and functionalities of the alleles in question. Furthermore, the observation that in short regions of the chromosome haplotypes tend to cluster together means that implementation of a Markov model (in essence, the form of a hereditary unit only depends on the current state of the unit, and not any of its past states) is often necessary to increase the accuracy of prediction genotypic data. It is this process of using many statistical tests that enables us to achieve fine localisation of important genes and begin to investigate multi-locus genotypes (with each gene only having a minute phenotypic effect). Statistics has given us a basis to build upon Mendelian genetics, by considering the interaction between genes in an organism. In addition, a firm foundation in statistics allows us to implement logistical regression, where we can work backwards and identify specific genetic markers (and potentially covariates of them) by observing the phenotype.

Statistics also has a massive role to play in medical research. It is particularly important in clinical trials, which is the process of testing medical interventions, and assessing the outcomes of these treatments. The nature of medicine often means that lives are at stake regarding various treatments. Therefore, safety is paramount, and as well as stringent regulations adopted in a clinical trial, a confidence interval of 99% is required before any new advancement is applied to the wider public. The goal of a clinical trial is to develop a new treatment with a higher efficacy than the old treatment. Most trials are comparative and are con-

ducted on a double-blind basis so neither doctor nor patient knows what treatment they are receiving. The patients used must also be randomised. Both of these steps reduce the bias in the trial, making the statistics used, valid. Multicentre trials increase sample size to increase reliability and reproducibility. The use of continuous interval scaling allows us to gauge the extent to which a treatment works, allowing a more thorough comparison of treatments rather than the nominal scaling. This can be coupled with the Spearman's Rank test, which is a non-parametric coefficient of correlation of two variables (-1 and +1 being perfectly monotonic relationships and 0 being no relationship), to show what quantitative effect a treatment has on a patient with certain criteria. In this way, statistical analysis allows us to show significant effects in otherwise seemingly ambiguous data and make conclusions that apply to the population as a whole.

In conclusion, statistics has a huge role in medical and scientific research, especially in genetics. This is because the cold-blooded nature of numbers reduces discrepancy in the interpretation of results, making new findings such as new treatments and discoveries, more accessible in the scientific community. Statistics is the foundation of all research, validating and disproving many ideas. The use of advanced statistics has deepened our understanding of genetics to such a level, that 'personalised medicines' are now prescribed by physicians tailored exactly to the genetic makeup of their patient. This powerful integration of genetics and medicine would not be possible without the use of statistics to process the raw data obtained.

Rohan Yesudian, The King's School, Chester

Australian Epigenetics Alliance Conference
29 October – 1 November 2017, Brisbane, Australia

This conference was attended by world-leading scientists from across the globe in the field of Epigenetics. The main goal was to foster and strengthen collaboration and unite researchers in the epigenetics field. Being held in a different state of Australia every other year, the conference is attended by a large number of Australian researchers who are often under-represented at European conferences. As such, attending this conference in Brisbane was a unique opportunity for me to engage with these renowned scientists and build a wider network of contacts.

As part of my current research, I am dissecting the transcriptome of spermatogenic cell populations on a single-cell level. This approach not only enables us to generate a gene expression atlas capturing the complex differentiation process of spermatogenesis, but also allows us to dissect the dynamic mechanisms underlying meiotic silencing – a surveillance mechanism that results in the transcriptional shut-down of unpaired chromosomes during meiosis. For this purpose, we are using a trans-chromosomic mouse model that carries a copy of human chromosome 21, which presents a unique model of aneuploidy that, unlike other trisomic mouse models, is able to complete male meiosis. This allows us to study the consequences of aneuploidy during this developmental process. While concealed in bulk experiments, our single-cell approach revealed differences in the level of meiotic silencing resulting in mosaicism in gene expression between spermatocytes.

Given my interest in single-cell techniques and analysis, I greatly benefitted from a variety of talks given at this conference that covered the latest developments in this field. These included a technique allowing the simultaneous analysis of DNA methylation, RNA transcripts and nucleosome occupancy within a given cell, which could give extremely valuable insights if applied to my model system. Since I am at the end of my PhD and am exploring different areas of the epigenetics field to pursue for my postdoctoral research, this conference provided an excellent overview, covering the importance of epigenetic regulatory mechanisms on a molecular level but also during cancer, development and ageing.

I was given the opportunity to present my research in form of a poster, as well as a short 'lightning' talk to spark interest in my work just before the poster session and was awarded with a poster prize for my presentation. The scientific discussions that I engaged in were extremely useful and provided me with invaluable feedback that will help in shaping the future direction of my work.

I therefore greatly benefitted from attending this conference, not only by gaining important feedback on my current work, but also by exploring potential new research avenues and connecting with a wider range of the epigenetics community. I am extremely grateful to the Genetics Society as well as the Galton Institute who jointly awarded me with a Junior Scientist Travel Grant and enabled me to attend this conference.

Christina Ernst
Cancer Research UK, Cambridge Institute

International Union for the Scientific Study of Population
Seminar on Urban Health Transformations

Organised by the IUSSP Historical Demography Panel & The Cambridge Group for the History of Population and Social Structure, July 2017 in Cambridge

Health in urban areas has played a major role in determining trajectories of demographic growth, economic success and individual and community well-being across time. However the relationship between health and urban space has not been constant over either time or place. Before the early twentieth century, towns and cities suffered a probably universal urban mortality penalty, and in some periods acted as 'demographic sinks', characterized by high death rates largely due to air and water-borne infections. The improvement of urban environments, together with the development of better preventive and curative medical services, which tend to be based in cities, means that urban areas today have lower mortality than surrounding areas. Although the decline of mortality in urban areas has been studied, there is little consensus about how urban spaces were transformed from unhealthy to healthy places. Such changes are unlikely to have happened at the same time or stage of industrial, economic or infrastructural development everywhere, but it has not been established whether there are any key developments which are necessary or sufficient for such transformations to occur. Attempts have been made to link declines in mortality to the introduction of sanitation and water supply, but with mixed success. The roles of housing, street paving, air pollution, and animal-keeping in fostering a hostile disease environment have been addressed less often. Municipal governance and institutions have been linked variously to poorer and to better health. How migration contributes to observed mortality rates is also poorly understood: migrants seeking work or a better life are often selected for better health, but may lack immunities to specific urban diseases. Chronic conditions such as tuberculosis may be linked to return or health-seeking migration, and such factors make it hard to disentangle the ways that migration, as other possible influences, might be linked to health outcomes. This meeting brought together researchers who addressed these, and related topics for different areas of the world from Cambodia to Copenhagen, and from the nineteenth century to the present.

The meeting started with two papers addressing the nature of the urban setting and how that affected mortality among older adults. Danan Gu, from the United Nations Population Division (USA), considered exposure to urban life and mortality in older cohorts in China, focussing on migrant status and the length of exposure to the urban setting. Matthias Voigt, from the Spanish National Research Council, examined how the effect of sociodemographic characteristics on health outcomes varies by levels of urbanicity in Andalucia, Spain, where urbanicity was defined using an index based on a variety of factors including population density, health and living conditions, occupations and perceptions. Both papers concluded that despite lower mortality in urban areas overall, once socio-economic conditions and other factors were fully controlled there was a higher risk of mortality in urban areas. A lively discussion ensued regarding what it is to be urban and the possibility of the renewal of an urban mortality penalty.

The next session continued with the theme of migration and mortality, specifically the notion that rural to urban migrants will 'come home to die' if they contract a chronic disease such as AIDS or tuberculosis. The two papers were from very different data settings and time periods: Carren Ginsburg (University of the Witwatersrand) focussed on Kenya and South Africa using recent Health and Demographic Surveillance Data, whereas that of Eilidh Garrett (University of Essex) used linked census and civil registration data for nineteenth century Scotland. Nevertheless they had many parallels, not only in their strong support for the 'returning home to die' hypothesis, but also regarding the importance of gender differences in behaviour – whether that was health seeking behaviour or the propensity to migrate in the first place. Both papers also commented on implications for broad theories: Ginsburg argued that the difference in risk by migrant status is better explained by structural socio-economic issues rather than by the stage of the epidemiologic transition; Garrett's demonstration that the risk of dying from TB was higher for young males than females working in textile and non-manual occupations casts doubt on the link between nutrition and TB which is at the heart of Thomas McKeown's theory of mortality decline.

During the conference there were two sessions relating to the development of sanitation and water supply in nineteenth century cities, and the first of these sessions was focussed on the UK. In her paper Romola Davenport (University of Cambridge) argued that the common use of infant mortality as an

indicator of water quality is misguided as breast-fed infants were more vulnerable to diseases (such as diarrhoea) spread by food or insect vectors. However cholera, highly dependent on water for transmission, makes an appropriate 'sanitary test'. Her results suggest that cholera was not very sensitive to piped water, and she concluded that the introduction of piped water did not necessarily result in a higher quality water supply. The challenges of how to measure sanitary investment and the development of water supply and sanitation in Britain were addressed by Bernard Harris (University of Strathclyde). He described a newly collected data source on loans issued by Parliament under Local Acts which shows more complete coverage of money available for such investment at a local level between 1817 and 1914, broadly confirming the previously incomplete picture of increasing sanitary effort towards the end of the nineteenth century. Andrew Hinde (University of Southampton) then used national and local mortality series, together with local case studies, to demonstrate that even with more accurate data, it is difficult to demonstrate a relationship between investment and mortality change as loans were often the culmination of a longer discussion and of less capital intensive initiatives. He argued that loans might be best understood as an indicator of the strength of concern related to public health.

The two papers in the next session examined health within urban areas. Ankita Shukla (Population Council, India) investigated the effect on mortality of having higher or lower levels of deprivation than surrounding areas. She found strong effects for districts in India, and considered the role of infrastructure and psychosocial effects. In her paper Bárbara Revuelta-Eugercios (University of Copenhagen) addressed the challenges of interpreting measurements of birth-weight from historical populations. Most historic birth-weight measurements derive from hospital deliveries which are often highly selected, so these new data from a large number of home deliveries are an important data source. Preliminary analysis indicates that the percentage of low birth-weight infants in Copenhagen in 1927 was not significantly larger than today, and that maternal age and parity were the most important determinants of low birth-weight.

The second day started with four papers on maternal health. The first two papers addressed access to health care in low and middle income countries. Myriam de Loenzien (CEPED-IRD, France) addressed urbanization as a determinant of increasing caesarean delivery in Vietnam and Cambodia, where caesarean rates are twice as high in urban areas compared to rural. Her analysis

showed that health infrastructure, socialisation processes, and perceptions of the baby's size were important determinants of the different rates. Eric Koba (Institut de Formation et de Recherche Démographiques, Cameroon) and Donita Nshani Tata (The University of Yaoundé I, Cameroon) addressed access to maternal health care in Senegal, finding that take-up of care was higher in urban areas and among higher socio-economic and more educated groups. Together these two papers suggested that both availability and individual choice are important determinants of maternal health behaviour. The final two papers in this session addressed maternal mortality in nineteenth century Europe. Angelique Janssens (Radboud University, Netherlands) examined the spatial patterns of maternal mortality in the Netherlands. She showed a correlation between maternal mortality and female tuberculosis mortality, particularly in rural areas, and argued that this lent strength to a picture of chronic female undernourishment. Alice Reid (University of Cambridge) compared different ways of measuring maternal mortality to demonstrate that many maternal deaths could not be identified using the recorded cause of death. She showed that under-recording was much greater in urban than rural areas where more deaths were certified by doctors who may have given a narrowly correct cause which omitted to mention a recent birth. Accounting for this reversed the rural-urban gradient in maternal mortality and she cautioned against urban-rural comparisons in causes of death where medical certification of death was highly variable.

In our second session on sanitation and health in the past Martin Önnerfors (Lund University, Sweden) used a detailed dataset relating to water processing and sewerage in Swedish cities to show that the introduction of water processing and sewage provision was strongly correlated with decline in water-borne (but not airborne) diseases, that the type of processing was important, and that it was better to have both water and sewerage than just one. Michael de Looper (Australian National University) provided a focussed look at the political and practical forces behind the trajectory of water and sewerage provision in Sydney, Australia, demonstrating a clear link between the introduction of clean water and the decrease in mortality from typhoid. Diego Ramiro Fariñas of the Spanish National Research Council also focussed on a single city, Madrid, in the early twentieth century. His detailed dataset enabled street level water and sewerage provision to be linked to individual mortality risks. He found that infant and child mortality were much more strongly dependent on water supply than on sewerage, and that there were also very persistent intra-

urban inequalities in the speed and extent of service provision.

Our final set of papers was focused on illness. Helene Castenbrandt (University of Gothenburg) analysed the records of two sickness funds for the city of Gothenburg in the early twentieth century. The fact that one of these was a women's fund provides a rare opportunity to compare men's and women's sickness episodes, and she found that women tended to have fewer, but longer episodes than men. It is as yet unclear whether this was due to selection of women into paid work, gendered health seeking behaviour, or some other reason. Teke Johnson Takwa of the Central Bureau for the Census, Cameroon, examined rural-urban differentials in anaemia prevalence among children in Cameroon. Anaemia prevalence was lower in urban areas, where it was linked to retarded growth, household living standards and the source of drinking water, among other factors also influential in rural areas. He argued that increasing urbanisation is likely to contribute to a future decline in anaemia, but only if urban growth is also accompanied by improving water supply and living standards. The final paper was presented by Grazyna Liczbinska of Adam Mickiewicz University, Poland, who used death registers for four Polish cities in the 19th and early 20th centuries to reflect on health and mortality. She showed that infant mortality was higher in industrial cities and before the introduction of sewerage.

To conclude the seminar our guest discussant, Tim Dyson of London School of Economics, provided an insightful and (mainly) sympathetic closing commentary. During this he drew together some of the major themes of the seminar and raised some broader issues. These included the perennial problem of how to define 'urban' which challenges both historical and contemporary demographers, and the fact that the nature of the urban-rural contrast may change over time and with the level of urbanisation. Nevertheless, he noted that many of the issues covered by both historical and contemporary demographers are the same. He felt that this seminar brought these out well and can be seen as a *prima facie* case for the establishment of a new IUSSP panel on the urban sector.

We are very grateful to our various sponsors including the Galton Institute and to Sophy Arulanantham of the Cambridge Group for the History of Population and Social Structure, for helping the conference to run so smoothly.

Romola Davenport
University of Cambridge

LETTER TO THE EDITOR

Dear Sir,

Recently it has been suggested that longevity and high intelligence are linked. Attempting to show this in an extended family was not successful.

108 people living their entire lives within the 20th century had been tested for many things including IQ by the author. The results were remarkably stable across the age spectrum.

IQ ranged from 69 to 158 (standardised) with 150 representing the top 1%. Socially, they ranged from small farmers to university lecturers, were even balanced by gender and died between the ages of 16 and 92. Accidental death has been included. The results were as follows:

Death age/IQ

IQ	< 90	91-100	101-110	111-120	121-130	131-140	141-150	151+
Mean death age	76.4	58.3	72.2	73.1	78.8	78.6	72.0	71.9
Age range	69-89	16-90	56-83	51-92	47-91	57-92	41-88	55-83
Number	13	6	5	5	46	10	8	8

The group with the low age score includes a young man with a sports accident and the low numbers are reflected in the result. Otherwise the mid 70's seem to be the expected age of dissolution.

Patrick James
Swallowcliffe, Salisbury

BOOK REVIEW

A.N. Wilson: Charles Darwin - Victorian Mythmaker
Pub. John Murray, pp.438

I should have been concerned as soon as I read the title.

There have, over the years, been a huge number of biographies of Charles Darwin. Some have been akin to hero worship, others decidedly less so. None, however can have been received quite like A.N. Wilson's latest effort. Reviews have included comments such as:

> "It is a puzzle how anyone could get Darwin and his science so completely wrong."
> "How could someone so unqualified to write about Darwin and today's science of evolution dare make such brazenly overconfident claims, sweeping away decades of scholarship?"

Andrew Norman Wilson is a biographer, novelist and newspaper columnist with a penchant for historical biographies. He is very popular in some circles and his 2014 work *Victoria: a Life* was well received. However, he is not a scientist and has no scientific training. It shows.

On the plus side, it is extremely well written and, it would seem, well researched. The early chapters on Darwin's time in Edinburgh and Cambridge are particularly revealing and he paints an intriguing picture of the Darwin-Wedgwood dynasty. This comes as no surprise. His family had close ties with the Wedgwood pottery company, his father having been managing director, and he has covered this ground before in previous books. Nevertheless, right from the start, one can hear the steady drip, drip of character assassination.

Some of his comments range from just irritating to bizarre. For example, he claims that Charles Darwin always played down the influence of his grandfather (Erasmus Darwin) so as to make evolution HIS theory and calls the younger Darwin "temperamentally allergic to controversy". He repeatedly mocks Darwin's use of the term "our dear old mother" when speaking to his children about his devoted wife, Emma. He says that all his books were based on "mythic con-

tradiction". And he even questions Darwin's sexuality during his voyage on the *Beagle*. Poor Captain FitzRoy doesn't escape the verbal assault either and is painted as a manic depressive, obsessed with his health whose captaincy was that "of a martinet". As for Darwin's ill health, he receives little sympathy, describing his symptoms as "psychosomatic" and mostly "stress-related". You soon feel that you're reading about a rather pathetic hypochondriac rather than one of history's greatest scientists.

However, things get much worse once Wilson moves on to the 'science'. There are the basic errors:

- Viruses have cells.
- Mendel's theory is lethal to Darwinism.
- The science of New Genetics delivered its death blow to Darwin's theory.

Incidentally, this 'New Genetics" to which Wilson keeps referring seems to end, as far as he is concerned, with Watson and Crick in 1953. The last 60 plus years of phylogenetic research appears to have had little influence on Wilson. He makes the schoolboy error of confusing selective fitness with physical fitness and constantly describes the "struggle for existence" in terms of physical fighting. He believes that 'selfish genes' produce selfishness and at one point seems to confuse Darwinism with Lamarckism. He also claims that there is no fossil evidence to support Darwin. And the inevitable old chestnut: "evolution is just a theory".

He saves his most outrageous statements for the final chapters on Darwin's impact. Curiously, he compares neo-Darwinism with Kipling's "Just So Stories" and Potter's "Tale of Peter Rabbit". He states that Darwin's observation that in most species "many more offspring are produced than can possibly survive" is "plainly not true". Inevitably, he asserts that the Third Reich laws of racial hygiene were all based on "bogus Victorian science much of which had started life in the gentle setting of Darwin's study at Down House". In the end, he appears to lay the blame for most of mankind's ills at the door of Darwin and his supporters.

This is a book that sets out to antagonise and annoy. In that respect, it is very successful.

My advice would be to catalogue it under fiction, not biography.

Robert Johnston

THE GALTON INSTITUTE

Conference 2018

Genome Editing

To be held 31 October, 2018 in
the Wellcome Trust Lecture Hall at the Royal Society,
Carlton House Terrace, London

Speakers and topics:

Professor Robin Lovell-Badge, FRS
The hows and whys of genome editing

Dr Kathy Niakan
Exploring early human development using CRISPR-Cas9

Professor Austin Burt
Manipulating mosquitoes for malaria control

Professor Daniel Voytas
Developing crops for sustainable agriculture and food security

Professor Richard Ashcroft
Societal considerations on genome editing

Professor Jennifer Doudna, FRS
The Galton Lecture 2018: *Genome Editing - history and future*

Professor Emma Morris
Exploring and treating human disease with genome editing

****

Admission is free, but strictly by ticket available from:
The General Secretary at: executiveoffice@galtoninstitute.org.uk
or www.eventbrite.co.uk