ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Royal Statistical Society
Statistics in Society
Series A

Professor Deborah Ashby

J. R. Statist. Soc. A (2019)

Pigeonholes and mustard seeds: growing capacity to use data for society

Deborah Ashby

Imperial College London, UK

[*The address of the President delivered to* The Royal Statistical Society *on Wednesday, June 26th, 2019*]

Summary. The Royal Statistical Society was founded to address social problems 'through the collection and classification of facts', leading to many developments in the collection of data, the development of methods for analysing them and the development of statistics as a profession. Nearly 200 years later an explosion in computational power has led, in turn, to an explosion in data. We outline the challenges and the actions needed to exploit those data for the public good, and to address the step change in statistical skills and capacity development necessary to enable our vision of a world where data are at the heart of understanding and decision making.

Keywords: Communication; Design principles; Statistical capacity building; Statistical skills; Team science

1. Origins

In the 1830s, interested groups of people were meeting regularly to discuss statistics, and in 1834 the Statistical Society of London was founded (Rosenbaum, 1984). In 1887, this became the Royal Statistical Society and the Royal Charter (Queen Victoria, 1887) set out

> 'Whereas Our Right trusty and entirely beloved cousin, Henry, Third Marquess of Lansdowne, [. . .], Charles Babbage [. . . and others . . .] did, in the year of One thousand eight hundred and thirty-four, establish a Society *to collect, arrange, digest and publish facts, illustrating the condition and prospects of society in its material, social, and moral relations; these facts being for the most part arranged in tabular forms and in accordance with the principles of the numerical method*, and the same Society is now called or known by the name of "The Statistical Society".
>
> 'And Whereas it has been represented to Us that the same Society has, since its establishment, sedulously pursued such its proposed objects, and *by its publications (including those of its transactions), and by promoting the discussion of legislative and other public measures from the statistical point of view, has greatly contributed to the progress of statistical and economical science.*
>
> 'And Whereas distinguished individuals in foreign countries, as well as many eminent British subjects, have availed themselves of the facilities offered by the same Society for *communicating important information largely extending statistical knowledge*; and *the general interest now felt in Statistics has been greatly promoted and fostered* by this Society.'

The Charter then goes on to address how a Society should be properly run, including the injunction 'There shall be a President . . .'. As the current holder of that office, I have the opportunity tonight to reflect on our past and to look to our future.

The language in the Charter may seem a little flowery, but those first three paragraphs set out that, from its earliest days, the Society has been about using data for the public good, developing

Address for correspondence: Deborah Ashby, School of Public Health, Imperial College London, Medical School Building, Norfolk Place, London, W2 1PG, UK.
E-mail: deborah.ashby@imperial.ac.uk

 0964–1998/19/182000

statistical science and building statistical capacity. I want to revisit what that meant then, and to think what that means in the 21st century.

The decade in which the Society was formed was a pivotal one for official statistics. In 1836, the General Registrar Office (GRO) was set up following the Births and Deaths Registration Act 1836 and began civil registrations of births and deaths the following year. These included data on occupation, which allowed, for example, calculation of occupational mortality rates. In 1837 the GRO also became responsible for the UK census of 1841, which was the first of the modern censuses, recording demographic details of individuals within households rather than just counting them. This allowed demographic analysis of the UK population, and subsequent censuses have built on this model. A key figure in both of these started as a temporary recruit to the GRO: the young William Farr. He was kept on at the GRO, and joined the Society in 1839, taking on many roles, including President in 1871. He published a detailed description of how to construct life tables for use in public health, building on Halley's work on their use for life annuities (Farr, 1859). In the same era, another Fellow, Charles Babbage, listed above as one of the sponsors on the Royal Charter, had been addressing the thorny question of automating calculations, developing a difference engine, and then an analytical engine. The British Government declined to fund its construction, although Farr did later order a Swedish version, for the GRO to use to calculate its *British Life Table*, using the data that they were now collecting, and an appendix to the resulting publication mentions some of the problems that they had had using it (Farr, 1864). This is an early example of the 'big data' of the day necessitating methodological computational innovation and leading to interesting debates about the role of the machine *vis-à-vis* that of the human 'calculators', as they were known. It exemplifies why our founding fathers were grappling with those three themes of using data for the public good, developing statistical science and building statistical capacity and also illustrates just how intertwined they were.

William Guy gave an Honorary Secretary's address (Guy, 1865) with the fabulous title

> 'On the original and acquired meaning of the term "Statistics", and on the proper functions of a Statistical Society, also on the question of whether there be a Science of Statistics; and, if so, what are its nature and objects and what is its relation to political economy and "Social Science"'

in which he discussed in a very scholarly fashion what is statistics. He proposed a first meaning (derived from 'state') that we might now understand as 'data', although he discouraged the improper use of the word statistics as a mere synonym for the collection of facts, irrespective of the use to which they are put. The second meaning, after a learned discussion of what is meant by 'science', is as 'Science of Statistics'. He had a comprehensive view of which areas the Society should be concerned with, including

> 'education, crime, industry, health, wealth, manufacture, commerce, special branches of industry and production &c'.

By contrast his (second) inaugural address as President started off more as a report to Council, all about building moves and membership, in particular who should be a Fellow and the need for growth (Guy, 1874). He aspired that the Society 'ought to count its members by the thousands not the hundreds', which strikes me as serious ambition to build statistical capacity. He finished by considering broader issues including commemorating John Howard's work on prisons, and 'offer[s] some remarks on the Conditions and Prospects of Society' which are wide ranging, and grappled, among other issues, with the interactions between poverty, 'vagrancy', 'imbeciles' and prisons. Guy's addresses illustrate eloquently how the Society was grappling with its founding themes.

Another towering figure was Charles Booth, from a wealthy Liverpool family, and well connected through family and marriage. As well as running a successful leather business and shipping line, he was very concerned about poverty and its measurement, finding data from the existing censuses inadequate for these purposes. Booth was awarded the first Guy Gold Medal, in honour of William Guy, by the Royal Statistical Society in 1892, in recognition of his achievements. He was also elected President in 1892 and, having advised on the 1891 census, presented its first results for the London population in his Presidential address of 1893 (Booth, 1893).

He had embarked on a massive data collection exercise, aided by his cousin through marriage Beatrice Potter (later Webb), which showed that 35% of Londoners were living in abject poverty. He produced a series of stunning poverty maps of London, which powerfully communicated his findings. He used this, and other work read to the Society in 1894 (Booth, 1894), to argue for the introduction of old age pensions. His work demonstrates his absolute commitment to gathering the appropriate data to learn about social conditions via the census and other means, the development of some powerful graphical methods in the pursuit of these aims and the use of these findings to argue for action.

Florence Nightingale became the Society's first female Fellow in 1858, with lead proposer William Farr. Known as 'the Lady with the Lamp', she is also recognized as a pioneering and passionate statistician (McDonald, 1998). Having persuaded her father to let her be educated in mathematics, she trained as a nurse and in 1854 went to Turkey as part of the Crimean War campaign, but she soon became concerned at the high death rate, and lack of decent records, which she put to rights. She worked with William Farr to analyse the data from the Crimean campaign, which she then used to procure royal support for a royal commission on the health of the army. She later turned her attention to the record keeping in London's hospitals, doing far more to improve health through these initiatives than she did directly via her nursing. She developed some stunning visualization methods, including the Coxcombe chart, to communicate complex information.

What is perhaps less widely recognized is Nightingale's passion for education, not only for herself, but also for others. Attewell (1998) has documented this clearly. She had taught at a Ragged School in Westminster, which gave her insights into the effects of poverty, and later used her influence to champion the education of British soldiers and army doctors. She set up the Nightingale School of Nursing at St Thomas's Hospital. She insisted that each probationer had a private room for study and reflection, but also had guiding principles that were practical and geared towards training nurses and then using them to spread her system widely. Her view on education, echoing her practical philosophy, was summarized in a quote that she took from an address delivered at the Universities of St Andrew's and Glasgow '[...] education is to teach men not to know, but to do'. McDonald (2001) highlights that Nightingale was concerned that, although data were routinely collected by government departments, there was a failure to use them in decision making. In a letter to Benjamin Jowett, Nightingale (1891) says

> 'Our chief point was that the enormous amount of statistics at this moment available at their disposal (or in their pigeon holes which means not at their disposal) is almost absolutely useless. Why? Because the Cabinet ministers... their subordinates, the large majority of whom have received a university education, have received no education whatever on the point upon which all legislation and all administration must—to be progressive and not vibratory—ultimately be based. We do not want a great arithmetical law; we want to know what we are doing in things which must be tested by results.'

Her proposed solution, together with Benjamin Jowett and Francis Galton, was to introduce statistics into studies at the University of Oxford by setting up the Professorship of Applied Statistics to address the need for statistics relating to education, penology, workhouses and

India, although, in the event, this did not happen. I suspect that her aspiration was just too practically focused. She was especially critical of the education that was available to women in her age, as well as the limited expectations that society had of their role (Attewell, 1998). She was primarily interested in action and said, in the context of sanitary reform and promotion of sanitary education in India, which was going slowly,

> 'I never lose an opportunity of urging a practical beginning, however small, for it is wonderful how often in such matters the mustard-seed germinates and roots itself'

(Cook, 1913).

To a very large extent, the early Fellows were concerned with what we now term observational data, whether collected routinely, or for a specific purpose. However, the idea of experimentation, which had its origins in the physical sciences, was beginning to evolve. At the Rothamsted Experimental Station (now Rothamsted Research), near Harpenden in Hertfordshire, a series of long-term agricultural experiments was instituted, known as the 'long term experiments' or 'classical experiments' (Perryman *et al.*, 2018; Poulton, 1995, 1996). The original purpose was to investigate the effects on crop yields of inorganic compounds containing nitrogen, phosphorus, potassium, sodium and magnesium. As knowledge emerged, experimental designs evolved to incorporate other questions. This provided fertile ground, not just for crops but for statistical development as well. Fisher was appointed as the first statistician at Rothamsted in 1919 to see whether he could do anything with the mass of data that they were accruing. This led to the development of important ideas for design as well analysis of data. His initial focus was on providing reliable estimates of the errors in field trials, and thus providing a basis for assessing whether treatment effects are real. The design concepts were not all new ideas, but he developed a coherent theory, analysis of variance, for analysing trial results incorporating these concepts (Yates, 1963).

David Finney played a leading role in statistics in agriculture in Scotland and read a paper to the Society on the statistician's role in the planning of field experiments, advocating the close involvement of the statistician in every stage (Finney, 1956).

Although there was a strong tradition of experimentation in agriculture, as far as human health went, the emphases in the early years of the Society were demography and the social determinants of health. Armitage (1983) described in his Presidential address that statisticians were more interested in counting and classifying deaths than in procedures of medical practice, although there had been exceptions such as studies into the efficacy of smallpox vaccination in the 18th century, and Lind's testing of six strategies of treating scurvy (Dunn, 1997). Armitage's writing in the early 1980s charted the development of the statistical approach in clinical medicine, including some discussion of the principles of experimentation, and their fruition. He dated the modern era as starting with the first properly randomized trial of streptomycin for the treatment of pulmonary tuberculosis (Medical Research Council Streptomycin in Tuberculosis Trials Committee, 1948), and he credited Austin Bradford Hill with playing a major role by introducing into clinical medicine the principle of experimentation that had previously been enunciated by Fisher. Armitage showed how the need to collect appropriate data stimulated more developments in statistical science and, by realizing the burgeoning need, also played a major part in capacity development of this area. When he succeeded Bradford Hill as Professor of Medical Statistics at the London School of Hygiene and Tropical Medicine, he planted some mustard seed by starting a Master of Science course in medical statistics, which graduated its first students in 1969. That seed took root. The Master's degree has just had its 50th-birthday celebrations and its hundreds of graduates, including me, are a testament to Peter Armitage's foresight.

We turn now to the present day to see how the themes of using data for the public good, developing statistical science and building statistical capacity are playing out currently.

2. Current challenges

The Royal Statistical Society's vision, as currently stated, is a world with data at the heart of understanding and decision making, and our four strategic objectives are as follows:

(a) for statistics to be used effectively in the public interest;
(b) for society to be more statistically literate;
(c) to develop the skills of the statistical profession;
(d) to strengthen the discipline of statistics.

Although these are expressed more succinctly than in the Royal Charter, they draw strongly on its vision, but with a stronger emphasis on the need for wider statistical literacy.

2.1. Using data for the public good

Our founding fathers made the case for data and focused on how to collect them. 'Big data' has been a recent buzz word, but what counts as big is relative. For William Farr, the mortality data and, for Charles Booth, the 1891 census data for London must have seemed quite big. Data collection in many spheres has become more routine, and with current technology we can do it at unprecedented scale, so new challenges are emerging.

A theme that cuts across all areas of statistics is that of open data, and how to balance interests such as privacy of the participants, and legitimate interests of those who have invested years of their careers in designing and collecting data, or company resources in the expectation of making a return to shareholders, with those who wish to access and make use of a resource that they regard as publicly funded (whether directly or indirectly), and that could be more widely used for the public good.

Very different traditions have emerged in different areas. Census data include highly personal details. Booth had a footnote on confidentiality arrangements in his analysis of the 1891 census (Booth, 1893). The data are traditionally released after 100 years, when most people who were in the census are no longer alive. Data from trials, by contrast, have not conventionally been made available, although changing perceptions and campaigns including AllTrials (Brown, 2013) are challenging that, and pharmaceutical companies have committed to making at least results, and often raw data, available, through initiatives such as clinicalstudydatarequest.com (So and Knoppers, 2017). Many public sector research funders stipulate that data should be made available after the end of the study, although the logistics and resourcing of this are still a difficult area. Another dimension is whether there are any controls on the purpose for which data are used. The census data, once available, can be used for genealogy, social research into occupations and by anybody for anything. In my own area of public health, by contrast, the 'Clinical practice research datalink' contains anonymized general practice and linked medical records for 11.3 million patients (Herrett *et al.*, 2015). The data are made available, but there is a committee that scrutinizes proposals, and data are only given out against a clear scientific objective. Big collaborations of cohort studies such as the European Prospective Investigation into Cancer and Nutrition Consortium, with prospective data on 500000 individuals in 10 European countries, work in much the same way (Bingham and Riboli, 2004).

2.2. Developing statistical science

We have already noted Farr's attempts to use Babbage's ideas on an analytical machine to construct life tables (Farr, 1864), and recent censuses have been analysed by using computers. Computers have also facilitated making previous census data available for interrogation and analysis in databases and formats that would astound those who set them up more than 100 years ago. Increased availability of data has been enabled by, and stimulated the need for, developments in methods for handling them, including computational advances, and statistical methodology. Peter Diggle outlined a brief history of statistics and information technology in his Presidential address, showing the opportunities that the rise of data science gives to statistics (Diggle, 2015).

Many of Booth's maps are now available on line from the London School of Economics and Political Science (London School of Economics and Political Science, 2016). Work on mapping has continued, and modern analogues can be used to help to investigate environmental correlates of ill health down to census ward level (Hawkes, 2014). Since then, spatial statistical methodology has developed enormously. The British regional heart study (Shaper *et al.*, 1981) was set up to explore some of the causes of geographical variation in heart disease. Early analyses explored these by using routine data and methods, novel at the time, that tweaked the existing multiple-regression methods, by incorporating a measure of distance (Pocock *et al.*, 1981, 1982). Further work has evolved to be much more explicit about the spatial properties of the mapping and has been applied in diverse areas including forestry and terrorism (de Rivera *et al.*, 2018; Python *et al.*, 2019).

To exploit the potential of data fully, there has been a major investment in 22 centres by a collaboration of 10 funders in Health Data Research UK with the vision that every health and care interaction and research endeavour will be enhanced by access to large-scale data and advanced analytics.

Recent developments in methodology for exploiting data include artificial intelligence (AI) and machine learning (ML). A Royal Society report looked at the potential of ML, but also at public understanding and perceptions of it (Royal Society (Great Britain), 2017). The Turing Institute has been set up as a national institute of data science and AI, convening groups to address a wide range of issues across domains such as finance and economics, engineering, defence and security as well as health.

These initiatives largely focus on exploiting observational data. Since the early days at Rothamsted, the design of experiments has evolved and been taken up into other areas (Atkinson and Bailey, 2001). In the health arena, following the streptomycin trial (Medical Research Council Streptomycin in Tuberculosis Trials Committee, 1948), the use of the experimental method to establish the safety and efficacy of medicine and other treatments has become widespread. A compelling demonstration of the potential influence of trials came from children's leukaemia, where the 5-year survival rate increased from 37% in 1971–1973 to 66% in 1980–1982, thanks to the systematic use of a series of standardized protocols incorporating entry to the UK acute lymphoblastic leukaemia trials (Stiller and Draper, 1989). More generally, Lewis (1983) put down an important marker on trials and their statistical issues, and the development of newer approaches to trial design and analysis has also featured in the Society's journals, notably from a Bayesian perspective (Racine *et al.*, 1986; Spiegelhalter *et al.*, 1994; Senn, 2000).

One of the drivers for this has been the increased emphasis on rigorous methods for the evaluation of medicines following the introduction of medicines regulation in the 1960s in the UK, USA and elsewhere. As well as the development of methods, and the explosion in demand for statisticians in the pharmaceutical industry, a parallel need for expertise within regulatory authorities became evident. The Royal Statistical Society realized the importance of this and wrote a report that resulted in the first statistical appointment within the UK authority, the Medicines Control Agency (now the Medicines and Healthcare Products Regulatory Agency). The statistical team then grew and has had an important influence within the European Medicines Agency

(Lewis, 1996). As drug development and regulation evolve, new statistical challenges appear. Following the catastrophic reactions of patients who were tested with the new drug TG1412, there was a wide-ranging review of practice in this area and the Royal Statistical Society set up a working party specifically on statistical issues in first-in-man studies (Senn *et al.*, 2007).

Advances in trial methodology, such as the development of cluster-randomized trials, enable the testing of interventions best randomized at a group level (Eldridge *et al.*, 2008; Hemming *et al.*, 2017) whereas the emergence of adaptive trials makes efficient use of scarce resources, be they time, enough patients to go into the trial or cost of the trial (Bhatt and Mehta, 2016). Within regulatory authorities and pharmaceutical companies there is interest in more rigorous ways of making decisions about the licensing of medicines, drawing on well-understood, but less well-used, statistical principles, in combination with other skill sets (Hughes *et al.*, 2016).

A live issue, to which statisticians regularly contribute, is the debate about the relative merits of the experimental method *versus* observational data, or 'trials *versus* real world evidence' as it is often termed (Rawlins, 2008). The existence of a large amount of readily available data gives many hope that this will somehow replace the need for careful experimentation. In my view, these different sources of data are complementary, answering rather different questions, and statisticians have a pivotal role in ensuring that different types of data are used appropriately.

2.3. Building statistical capacity

Statistical capacity building is a recurrent theme in recent Presidential addresses. Adrian Smith focused on mathematics education in schools, David Hand discussed the statistical education from the next generation of statisticians through to the lay public, and Valerie Isham had a strong focus on education and the need for nurturing the discipline. John Pullinger regarded statistical education and literacy as essential for impact, and Peter Diggle noted 'We are what we teach' and emphazised especially the role of design (Smith, 1996; Hand, 2009; Isham, 2012; Pullinger, 2013; Diggle, 2015).

Within the Royal Statistical Society, the emergence of a vibrant Young Statisticians' Section has energized the Society and acted as a focus for growing the future generation. Our professional qualifications of Graduate Statistician and Chartered Statistician provide a benchmark, and continuing professional development is mandated for all Chartered Statisticians. The Society has developed a suite of training courses for statisticians and data scientists. There is major investment in training from the Research Councils, e.g. through the Engineering and Physical Sciences Research Council and UK Research and Innovation's recent allocation of Centres for Doctoral Training, and the National Institute for Health Research's investment in training including schemes that are designed to attract talented individuals to train in statistics and allied disciplines: skills critical to improving health and wealth in the UK. Both Health Data Research UK and the Turing Institute are devoting considerable resources to training and capacity development.

The Royal Statistical Society continues to champion the importance of statistics externally. Its training courses include foundation level courses including a free short on-line training course in basic statistics for journalists. Our 'Data manifesto' outlined a positive vision of how statistics can contribute to democracy, policy and prosperity. Many members wrote to their parliamentary candidates during the 2015 election, raising the profile of the manifesto and asking candidates to pledge to take statistical training from the Society if they were elected. Over 300 pledged to do so, of whom 55 were elected, a handful of whom then actually took up the training. We also have our 'Statisticians for society' *pro bono* work which has now connected statisticians to over 50 charitable projects. As part of this programme we also have our work with the African

Institute of Maths Sciences, sending Fellows to work on a voluntary basis in Africa, with their travel supported by a partnership with Taylor and Francis.

3. The Royal Statistical Society at 200

In 15 years' time the Society will celebrate its 200th anniversary. The areas of education, crime, industry, health, wealth, manufacture, commerce, special branches of industry and production, that Guy thought so important, continue to pose new challenges. We also face new challenges. Improvements in living conditions and healthcare have resulted in a much greater life expectancy so, while we celebrate that, we also need to address the consequences including how we might research and provide healthcare and social care for people with increasing multimorbidity (Whitty, 2017) and a pensions crisis, the scale of which Booth could not have foreseen when he argued for the introduction of old age pensions. An area that our founders did not consider is the environment. The award this year of the Guy Medal in Gold to Stephen Buckland and the Barnett Award for Environmental Statistics to Marian Scott highlight the long-standing contribution of statistics and statisticians in this area, which is becoming even more pressing with rapid climate change. A just released summary of a forthcoming major United Nations report is showing global data on loss of diversity and species loss with potentially catastrophic consequences. Faced with huge and complex issues such as these, what are the challenges and opportunities for the Royal Statistical Society to continue to pursue its vision in a world with data at the heart of understanding and decision making?

3.1. Using data for the public good

As a society, we already now have more data available than we have ever had. Advances in technology mean that increasingly more data are accruing, whether in health records, on mobile device applications, from our electronic payments for travel and day-to-day purchases. That offers huge potential for good, but also limitless opportunities for misinterpretation or misuse. My predecessor addressed the issue of trustworthiness (Spiegelhalter, 2017). We are in danger, to misquote the Rime of the Ancient Mariner, of having 'Data, data, everywhere, nor any stop to think'. Like salty seawater, data can either drown or sustain us, but we need to know how to use them. And, to push the analogy, whether we want them with or without saltiness depends on the use we need to make of them.

Each decennial census has evolved from the previous one, and we are just gearing up for the 2021 census. At one level this is a continuation and evolution of the census that we have had since 1801 (Office for National Statistics, 2018a) but, as it is the first to use on-line data collection (Office for National Statistics, 2018b), it will feel quite different to those participating. It gives great opportunity for public debate and dialogue about the value of data, as well as issues of confidentiality and trustworthiness.

Spatial analysis takes on new meaning. Recent developments mean that data generated on social media (such as searches for symptoms of influenza) have been used for the tracking of epidemics although this turned out to be problematic (Lazer *et al.*, 2014) and mobile phone data, which contain geographic locators, can be used, for example, to estimate the size of crowds (Botta *et al.*, 2015) and to study wildlife population abundance (Yuan *et al.*, 2017), with environmental sciences offering fresh challenges (Scott and Gemmell, 2012).

Health data sets are becoming more sophisticated, offering new opportunities not just for observational work, but also the possibility of using them to run pragmatic trials in routine clinical practice, obviating the need for specialist data collection (van Staa *et al.*, 2012).

The use of such data quickly leads to questions of privacy, and who should have access to data for uses beyond their primary purpose. There is often an implicit assumption that data are owned by an individual, but much data are 'owned' by more than one person. For example, genetic data at least implicitly give information about close relatives, and data on a visit to a general practitioner are relevant both to the patient and to the management of the practice. I prefer to think of data as a jointly owned resource—perhaps 'people's data' rather than 'your data' or 'my data'. It's a little like tax; we all contribute in various ways when we can, and we benefit when we need to. In areas such as health and education which happen largely in the public sphere, there may be some parallels with Richard Titmuss's analysis of blood donation and wider issues in his book *The Gift Relationship* which compares a non-market system based on altruism with a system based on for-profit enterprises, concluding that the former is far more effective (Titmuss, 1997). However, much data reside in the private sector. With increasing computerization, and widespread use of electronic means of payment and loyalty cards as well as social media, we are leaving data trails almost without realizing it. Those data, when collated and analysed, can give information that has a high value, e.g. on spending patterns. A further development is that it can then be used for targeted advertising based on previous spending or searches. This has taken on new levels of sophistication with the uses of data on networks of individuals to influence election campaigns. The Cambridge Analytica case has led to extensive public debate on the uses to which data can and should be put (Isaak and Hanna, 2018). That inevitably has led to debate about whether there is a need for regulation in the area and what the ethical principles are. A plethora of structures and organizations are emerging, including the Centre for Data Ethics and Innovation, which

> 'is an advisory body set up by Government and led by an independent board of expert members to investigate and advise on how we maximise the benefits of data-enabled technologies, including artificial intelligence'

and the Ada Lovelace Institute, of which the Royal Statistical Society is a founding partner, which describes itself as

> 'an independent research and deliberative body with a mission to ensure data and AI work for people and society'

and says that

> 'Ada will promote informed public understanding of the impact of AI and data-driven technologies on different groups in society'.

3.2. Developing statistical science

To be able to use the data that are available today for the public good, the continuing development of appropriate methods is critical.

The 2021 census is breaking new ground. Comparative work will enable calibration of older and newer data collection methods and plans to integrate with other data to produce and disseminate census outputs will set the foundations for producing up-to-date population statistics beyond the 2021 census (Office for National Statistics, 2018). The combination of methods development with modern data collection methods gives huge potential to use the census data for the good of society.

Doing excellent statistics is of limited use without communication (Bles *et al.*, 2019). Nightingale produced visualizations of her data to convey her messages to those who could act, and Booth produced his poverty maps. We now have the ability for users to interact with visualizations and other displays, to choose formats, to explore further. There are some great examples, e.g. on the 'Understanding uncertainty' website (`https://understandinguncertainty.org`)

(Winton programme for the public understanding of risk), and there are huge further opportunities to use computational power to interrogate and display data.

In trials, a particular area of development is that of adaptive designs, which aim to maximize the information from trials by allowing changes while the trial is running in response to accruing results. This might, for example, be the choice of dose (Mason *et al.*, 2017), or dropping certain patient groups. The key feature is that these should be prespecified, and the characteristics of the design well understood. The development of these studies has technical challenges in their design and logistic challenges in their execution, and they often need intensive computational methods. They also pose challenges to those regulating new medicines, as they stretch the normal paradigm of exploratory and then confirmatory studies (Center for Drug Evaluation and Research, 2018).

One variant of adaptive design is known as the platform trial. Whereas a conventional trial tests one treatment against a comparator for a particular patient group, platform trials work as protocols where every patient enters an appropriate arm for them and, over time, the treatment choices in each arm may vary in the light of accruing evidence or new therapies emerging (Parmar *et al.*, 2008). A well-established example is the 'Systemic therapy for advancing or metastatic prostate cancer' trial in prostate cancer (James *et al.*, 2009), which may more properly be thought of as a co-ordinated set of trials. From a health perspective, these are the successors to the childhood leukaemia protocols that led to such fundamental changes in survival, but, from a design perspective, they might be seen as the successors to the 'long term experiments' at Rothamsted. The potential of such systematic approaches to learning is great, but so is the need for appropriate methodological development as the paradigm shifts. Personalized medicine is another area raising big expectations and will pose commensurate methodological challenges, including the ground rules for licensing.

The use of the experimental method is well established in health but is slower to take root in other domains such as education, crime, justice and environment. Stella Cunliffe, our first female President, was sad to 'have left the field of experimentation' (Cunliffe, 1976) when she joined the Home Office. Despite some good exemplars, e.g. in the health of prisoners (Bird *et al.*, 2017), the culture change that happened in health still seems a long way off, so the Royal Statistical Society has work to do.

Perhaps the biggest methodological challenges come from the abundance of data now accruing. A naive view is that size trumps everything; a cynical view is that it has all the problems of normal data writ large. There are those who hope that greater availability of data somehow obviates the need for well-designed experiments. One set of challenges to those working in data science relates to computational methods for storing, combining and using the data, on a scale that takes us out of traditional statistical territory, and into that of computer scientists (Diggle, 2015). Even when that is sorted, analysis is not straightforward. The primary purposes of such data are often very different from the secondary uses that we wish to make of them, and the traditional statistical skills of really understanding the data are still highly pertinent. An area where the observational studies have been at odds with trial evidence is the role of postmenopausal hormone therapy on the subsequent risk of cardiovascular disease and breast cancer (Prentice *et al.*, 2005). Many have thought hard about the reasons for the differences (Vandenbroucke, 2009). Hernan and Robins (2016) showed how 'big data' can be used to emulate a target trial by using design-of-experiment principles. Working through their approach gives real insights about why. There is also emerging thinking that there is an important, more general, role for the application of design principles for the analysis of such data (Drovandi *et al.*, 2017), and this seems to me a very important area for future development, to exploit them best in a statistically principled manner.

3.3. Building statistical capacity

To meet the biggest challenges facing society we need increasing statistical and data science capacity in order fully to be able to use appropriate methods with the data that are available today for the public good.

Florence Nightingale tried to increase statistical capacity via endowing a Chair, and her ambition to provide statistical education seems more relevant than ever. Data now offer huge promise, and on a scale of which our predecessors could only have dreamt. As data are accrued in new ways, there are equally huge needs and opportunities for developing methodologies to exploit them. The greatest challenge that I see is in growing the capacity to do that. When I analysed Swedish birth records as an undergraduate project, and first heard about the spatial analyses being carried out on the regional heart study, part of the excitement for me was that the methods being developed and applied were not a huge step up from those I was learning about in my undergraduate courses. I felt that making a difference was within my grasp. Although methods and principles that are taught at undergraduate level are still a great basis, the gap between those and cutting-edge methods is now far larger. That raises real challenges in how we train future generations, but also in how we develop skills of those who trained a couple of decades ago or more.

As well as individuals needing a greater depth of training, there is an ever-burgeoning increase in numbers of people who are needed with these skills. Some of these will be in the traditional statistical mould, some more mathematical; others will be data scientists, but many more need these skills as part of their portfolio, either to carry out the work, or to be able to understand and use them intelligently as part of their work. The Royal Society have just produced a report on data science skills, with case-studies and models of good practice, demonstrating clearly that statistics matters, and citing current activities and further opportunities for the Royal Statistical Society (Blake, 2019).

Those who work in AI and ML need to understand statistical principles, but it is equally important that statisticians should learn more about AI and ML, both for their own sakes, and to be able to engage with those communities. Part of the solution is more flexible models of education and training. As well as traditional campus-based courses, there is an explosion in on-line learning, ranging from individual massive open on-line courses, which are often available free or at minimal costs, to full Master's courses. These are well suited to areas such as ours, and the development and take-up of massive open on-line courses on topics such as AI indicate the phenomenal demand. For some, these will replace traditional campus-based study, but for others they can be a source of accessible continuing professional development.

Many statisticians are used to working as part of teams. Some researchers, such as people working at the Large Hadron Collider (Giudice, 2012) in physics, now work in huge teams as do many genetics collaborations. The idea of team science is emerging with discussion, for example, on the implications for careers (Libby *et al.*, 2016; Saunders, 2017). I think of data science as an activity of teams rather than individuals, with people complementing each other. Different people may need different levels of various aspects at different career stages, but it should also be possible to transition between them when needed. That can have profound implications for how we think about education and training. I think of statistics as a language for communicating what we can learn from data. The Common European Framework of Reference for Languages (Council of Europe, 2018), may be a helpful analogy in planning training and continuing professional development for individuals and groups. Explicitly learning how to work in 'team mode' is also a priority, with communication being a critical skill.

In the 1887 Royal Charter, the language was that of 'communicating important information', which sounds rather one directional. Nightingale's focus was on politicians, but the public also needs a greater level of understanding to benefit fully from data and not to be misled by far-fetched claims. In areas such as health, the language with respect to research now is very much more about public involvement, and working in partnership with coproduction of research (Staniszewska *et al.*, 2018). The census offers opportunities both in its design, but also in how it is rolled out, and how results are then communicated. The public here is not a homogeneous body, but mediated through the media, charities and elected representatives as well as individuals themselves. It is easy to think of this as 'us' and 'them', but, rather like team science, it is more constructive to think in terms of a spectrum of users working together, and the Royal Statistical Society, with its traditionally open welcome to all who are interested in data, is well placed to facilitate that. It has been doing excellent work with some of these groups and now needs to think how to step that up further.

The Royal Statistical Society has long-standing interests in educational policy and a new Special Interest Group on Teaching Statistics. Florence Nightingale's criticism of the lack of education that was available to women is a forerunner of today's concerns with diversity. Stella Cunliffe declined to devote her Presidential address to the theme of the position of women (Cunliffe, 1976). However, we are gradually making progress on gender diversity within the Society, although there have been more Presidents named David than women Presidents, and the same is true for George or William. We are beginning to think about diversity of other personal characteristics but, to address today's challenges with data, we also need a wide diversity of skill sets and perspectives working together, while recognizing that there is unity in that diversity (Armitage, 1983; Green, 2003). The Society's Diversity Working Group and the new Special Interest Group in Women in Statistics and Data Science are spearheading our efforts here.

Education may now be readily available to women in the UK and many other countries, but there are still many societies where that is not so. And there are other inequities in access to education. Nightingale championed education in very practical ways. We need to ensure that statistical capacity development reaches widely; otherwise we are in danger of perpetrating further inequalities, not least because data and statistical literacy are increasingly essential to function fully in society, and we need all hands on deck, which is not helped by statistically disenfranchising large sectors of society.

In putting this together, I am aware of the breadth of approaches and application areas on which I have touched. All of the areas and examples that were described above have many statistical issues to address, increasing the need for a high level of statistical expertise, including the ability to work in highly multidisciplinary scientific and legal frameworks. Recent developments take the need for all the skills that are required to another level both in terms of depth and the numbers of people needed (Kingman, 2018). If we rise to that, the potential for using the data for the public good is enormous.

4. Pigeonholes and mustard seeds

Florence Nightingale's frustration with enormous amounts of statistics in pigeonholes that do not feed through to good decision making resonates today, although those pigeonholes are largely replaced by databases and e-mail folders. William Guy's ambition for the Society's membership to go from hundreds to thousands has been fulfilled. The need now is for millions, even billions, of statistically literate people. The Royal Statistical Society has led the way in growing capacity since its foundation and is well positioned to continue to play a leading role in partnership with others. To do this, we should aspire that our membership should be counted in their tens of thousands, at the least. This may seem an insurmountable challenge, but I am inspired by Florence Nightingale's urging of a practical beginning which, like mustard seed, geminates, roots itself and grows in ways that we might not expect. The flourishing Society that we have now is the result of sowing of seeds by our founders. It may seem unrecognizable from those early meetings, but at its heart, nearly 200 years later, it is still all about using data for the public good, developing statistical science and building statistical capacity. I am very proud to be its President. Let us keep sowing those mustard seeds.

Acknowledgements

I thank Guy Nason, Hetan Shah and Rosalind Smyth for insightful and constructive comments on an earlier draft.

References

Armitage, P. (1983) Trials and errors—the emergence of clinical statistics. *J. R. Statist. Soc.* A, **146**, 321–334.
Atkinson, A. C. and Bailey, R. A. (2001) One hundred years of the design of experiments on and off the pages of *Biometrika*. *Biometrika*, **88**, 53–97.
Attewell, A. (1998) Florence Nightingale (1820-1910). *Prospects*, **28**, 153–166.
Bhatt, D. L. and Mehta, C. (2016) Adaptive designs for clinical trials. *New Engl. J. Med.*, **375**, 65–74.
Bingham, S. and Riboli, E. (2004) Diet and cancer—the European prospective investigation into cancer and nutrition. *Nat. Rev. Cancer*, **4**, 206–215.
Bird, S. M., Strang, J., Ashby, D., Podmore, J., Robertson, J. R., Welch, S., Meade, A. M. and Parmar, M. K. (2017) External data required timely response by the Trial Steering-Data Monitoring Committee for the NALoxone InVEstigation (N-ALIVE) pilot trial. *Contemp. Clin. Trials Communs*, **5**, 100–106.
Blake, A. (2019) Dynamics of data science skills. *Report*. Royal Society, London.
van der Bles, A. M., van der Linden, S., Freeman, A. L. J., Mitchell, J., Galvao, A. B., Zaval, L. and Spiegelhalter, D. J. (2019) Communicating uncertainty about facts, numbers and science. *R. Soc. Open Sci.*, **6**, article 181870.
Booth, C. (1893) Life and labour of the people in London: first results of an inquiry based on the 1891 census. *J. R. Statist. Soc.*, **56**, 557–593.
Booth, C. (1894) Statistics of pauperism in old age. *J. R. Statist. Soc.*, **57**, 235–245.
Botta, F., Moat, H. S. and Preis, T. (2015) Quantifying crowd size with mobile phone and Twitter data. *R. Soc. Open Sci.*, **2**, article 150162.
Brown, T. (2013) It's time for Alltrials registered and reported. In *Cochrane Database of Systematic Reviews*.
Center for Drug Evaluation and Research Center for Biologics Evaluation and Research (2018) Adaptive designs for clinical trials of drugs and biologics guidance for industry. Center for Drug Evaluation and Research, Silver Spring.
Cook, E. T. (1913) *The Life of Florence Nightingale*, vol. II. London: Macmillan.
Council of Europe (2018) Council of Europe Common European Framework of Reference for Languages: learning, teaching, assessment. Council of Europe. (Available from `https://www.coe.int/en/web/common-european-framework-reference-languages`.)
Cunliffe, S. V. (1976) Interaction. *J. R. Statist. Soc.* A, **139**, 1–19.
Diggle, P. J. (2015) Statistics: a data science for the 21st century. *J. R. Statist. Soc.* A, **178**, 793–813.
Drovandi, C. C., Holmes, C., McGree, J. M., Mengersen, K., Richardson, S. and Ryan, E. G. (2017) Principles of experimental design for big data analysis. *Statist. Sci.*, **32**, 385–404.
Dunn, P. M. (1997) James Lind (1716-94) of Edinburgh and the treatment of scurvy. *Arch. Dis. Childhd*, **76**, F64–F65.
Eldridge, S., Ashby, D., Bennett, C., Wakelin, M. and Feder, G. (2008) Internal and external validity of cluster randomised trials: systematic review of recent trials. *Br. Med. J.*, **336**, 876–880.
Farr, W. (1864) *English Life Table: Tables of Lifetimes, Annuities, and Premiums*. London: Her Majesty's Stationery Office.
Farr, W. (1859) On the construction of life-tables, illustrated by a new life-table of the healthy districts of England. *Phil. Trans. R. Soc. Lond.*, **149**, 837–895.
Finney, D. J. (1956) The statistician and the planning of field experiments. *J. R. Statist. Soc.* A, **119**, 1–17.
Giudice, G. F. (2012) Big science and the Large Hadron Collider. *Phys. Perspect.*, **14**, 95–112.
Green, P. J. (2003) Diversities of gifts, but the same spirit. *Statistician*, **52**, 423–435.
Guy, W. A. (1865) On the original and acquired meaning of the term "Statistics," and on the proper functions of a Statistical Society: also on the question whether there be a Science of Statistics; and, if so, what are its nature and objects, and what is its relation to political economy and "Social Science". *J. Statist. Soc. Lond.*, **28**, 478–493.
Guy, W. A. (1874) Inaugural Address delivered at the Society's Rooms, Somerset House Terrace, King's College, London, on Tuesday, 17th November, 1874. *J. Statist. Soc. Lond.*, **37**, 411–436.

Hand, D. J. (2009) Modern statistics: the myth and the magic. *J. R. Statist. Soc.* A, **172**, 287–306.
Hawkes, N. (2014) Atlas aims to show possible environmental effects on UK health. *Br. Med. J.*, **348**, article g2948.
Hemming, K., Eldridge, S., Forbes, G., Weijer, C. and Taljaard, M. (2017) How to design efficient cluster randomised trials. *Br. Med. J.*, **358**, article j3064.
Hernan, M. A. and Robins, J. M. (2016) Using big data to emulate a target trial when a randomized trial is not available. *Am. J. Epidem.*, **183**, 758–764.
Herrett, E., Gallagher, A. M., Bhaskaran, K., Forbes, H., Mathur, R., van Staa, T. and Smeeth, L. (2015) Data resource profile: clinical practice research datalink (CPRD). *Int. J. Epidem.*, **44**, 827–836.
Hughes, D., Waddingham, E., Mt-Isa, S., Goginsky, A., Chan, E., Downey, G. F., Hallgreen, C. E., Hockley, K. S., Juhaeri, J., Lieftucht, A., Metcalf, M. A., Noel, R. A., Phillips, L. D., Ashby, D., Micaleff, A. and PROTECT Benefit Risk Group (2016) Recommendations for benefit–risk assessment methodologies and visual representations. *Pharmepidem. Drug Safty*, **25**, 251–262.
Isaak, J. and Hanna, M. J. (2018) User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, **51**, 56–59.
Isham, V. (2012) The evolving Society: united we stand. *J. R. Statist. Soc.* A, **175**, 315–335.
James, N. D., Sydes, M. R., Clarke, N. W., Mason, M. D., Dearnaley, D. P., Anderson, J., Popert, R. J., Sanders, K., Morgan, R. C., Stansfeld, J., Dwyer, J., Masters, J. and Parmar, M. K. (2009) Systemic therapy for advancing or metastatic prostate cancer (STAMPEDE): a multi-arm, multistage randomized controlled trial. *BJU Int.*, **103**, 464–469.
Kingman, J. O. F. (2018) Speech at Royal Society 'Research culture: changing expectations' Meet. (Available from `https://www.ukri.org/files/news/jk-speech-301018-pdf/`.)
Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014) The parable of Google flu: traps in big data analysis. *Science*, **343**, 1203–1205.
Lewis, J. A. (1983) Clinical trials: statistical developments of practical benefit to the pharmaceutical industry. *J. R. Statist. Soc.* A, **146**, 362–377.
Lewis, J. A. (1996) Statistics and statisticians in the regulation of medicines. *J. R. Statist. Soc.* A, **159**, 359–362.
Libby, A. M., Cornfield, D. N. and Abman, S. H. (2016) There is no "i" in team: new challenges for career development in the era of team science. *J. Pediatr.*, **177**, 4–5.
London School of Economics and Political Science (2016) Charles Booth's London: poverty maps and police notebooks. London School of Economics and Political Science, London. (Available from `https://booth.lse.ac.uk/`.)
Mason, A. J., Gonzalez-Maffe, J., Quinn, K., Doyle, N., Legg, K., Norsworthy, P., Trevelion, R., Winston, A. and Ashby, D. (2017) Developing a Bayesian adaptive design for a phase I clinical trial: a case study for a novel HIV treatment. *Statist. Med.*, **36**, 754–771.
McDonald, L. (1998) Florence Nightingale: Passionate statistician (with discussion). *J. Holist. Nursng*, **16**, 267–280.
McDonald, L. (2001) Florence Nightingale and the early origins of evidence-based nursing. *Evid. Basd Nursng*, **4**, 68–69.
Medical Research Council Streptomycin in Tuberculosis Trials Committee (1948) Streptomycin treatment of tuberculous meningitis. *Lancet*, **251**, 582–596.
Nightingale, F. (1891) Letter to Benjamin Jowett.
Office for National Statistics (2018a) Office for National Statistics Census Transformation Programme. Office for National Statistics, Newport. (Available from `https://www.ons.gov.uk/census/censustransformationprogramme`.)
Office for National Statistics (2018b) Help shape our future: the 2021 Census of Population and Housing in England and Wales. Her Majesty's Government, London.
Parmar, M. K., Barthel, F. M., Sydes, M., Langley, R., Kaplan, R., Eisenhauer, E., Brady, M., James, N., Bookman, M. A., Swart, A. M., Qian, W. and Royston, P. (2008) Speeding up the evaluation of new agents in cancer. *J. Natn. Cancer Inst.*, **100**, 1204–1214.
Perryman, S. A. M., Castells-Brooke, N. I. D., Glendining, M. J., Goulding, K. W. T., Hawkesford, M. J., Macdonald, A. J., Ostler, R. J., Poulton, P. R., Rawlings, C. J., Scott, T. and Verrier, P. J. (2018) The electronic Rothamsted Archive (e-RA), an online resource for data from the Rothamsted long-term experiments. *Sci. Data*, **5**, article 180072.
Pocock, S. J., Cook, D. G. and Beresford, S. A. A. (1981) Regression of area mortality rates on explanatory variables: what weighting is appropriate? *Appl. Statist.*, **30**, 286–295.
Pocock, S. J., Cook, D. G. and Shaper, A. G. (1982) Analysing geographic variation in cardiovascular mortality: methods and results. *J. R. Statist. Soc.* A, **145**, 313–329.
Poulton, P. R. (1995) The importance of long-term trials in understanding sustainable farming systems: the Rothamsted experience. *Aust. J. Exptl Agric.*, **35**, 825–834.
Poulton, P. R. (1996) The Rothamsted long-term experiments: are they still relevant? *Can. J. Plant Sci.*, **76**, 559–571.
Prentice, R. L., Langer, R., Stefanick, M. L., Howard, B. V., Pettinger, M., Anderson, G., Barad, D., Curb, J. D., Kotchen, J., Kuller, L., Limacher, M., Wactawski-Wende, J. and International Women's Health Initiative (2005)

Combined postmenopausal hormone therapy and cardiovascular disease: toward resolving the discrepancy between observational studies and the Women's Health Initiative clinical trial. *Am. J. Epidem.*, **162**, 404–414.
Pullinger, J. (2013) Statistics making an impact. *J. R. Statist. Soc.* A, **176**, 819–839.
Python, A., Illian, J. B., Jones-Todd, C. M. and Blangiardo, M. (2019) A Bayesian approach to modelling subnational spatial dynamics of worldwide non-state terrorism, 2010–2016. *J. R. Statist. Soc.* A, **182**, 323–344.
Queen Victoria (1887) Royal Statistical Society Copy of 1887 Royal Charter.
Racine, A., Grieve, A. P., Flühler, H. and Smith, A. F. M. (1986) Bayesian methods in practice: experiences in the pharmaceutical industry. *Appl. Statist.*, **35**, 93–120.
Rawlins, M. (2008) *De testimonio*: on the evidence for decisions about the use of therapeutic interventions. *Lancet*, **372**, 2152–2161.
de Rivera, O. R., Lopez-Quilez, A. and Blangiardo, M. (2018) Assessing the spatial and spatio-temporal distribution of forest species via Bayesian hierarchical modeling. *Forests*, **9**, article 573.
Rosenbaum, S. (1984) The growth of the Royal Statistical Society. *J. R. Statist. Soc.* A, **147**, 375–388.
Royal Society (Great Britain) (2017) *Machine Learning: the Power and Promise of Computers that Learn by Example*. London: Royal Society.
Saunders, P. (2017) Supporting researchers in an era of team science. *Lancet*, **389**, suppl. 1, S10 –S12.
Scott, E. M. and Gemmell, J. C. (2012) Spatial statistics—a watery business. *Spatl Statist. Neth.*, **1**, 121–132.
Senn, S. (2000) Consensus and controversy in pharmaceutical statistics (with discussion). *Statistician*, **49**, 135–176.
Senn, S., Amin, D., Bailey, R. A., Bird, S. M., Bogacka, B., Colman, P., Garrett, A., Grieve, A. and Lachmann, P. (2007) Statistical issues in first-in-man studies. *J. R. Statist. Soc.* A, **170**, 517–579.
Shaper, A. G., Pocock, S. J., Walker, M., Cohen, N. M., Wale, C. J. and Thomson, A. G. (1981) British Regional Heart-Study—cardiovascular risk-factors in middle-aged men in 24 towns. *Br. Med. J.*, **283**, 179–186.
Smith, A. F. M. (1996) Mad cows and ecstasy: chance and choice in an evidence-based society. *J. R. Statist. Soc.* A, **159**, 367–381.
So, D. and Knoppers, B. M. (2017) Ethics approval in applications for openaccess clinical trial data: an analysis of researcher statements to clinicalstudydatarequest.com. *PLOS One*, **12**, article e0184491.
Spiegelhalter, D. (2017) Trust in numbers. *J. R. Statist. Soc.* A, **180**, 949–965.
Spiegelhalter, D. J., Freedman, L. S. and Parmar, M. K. B. (1994) Bayesian approaches to randomized trials. *J. R. Statist. Soc.* A, **157**, 357–387.
van Staa, T. P., Goldacre, B., Gulliford, M., Cassell, J., Pirmohamed, M., Taweel, A., Delaney, B. and Smeeth, L. (2012) Pragmatic randomised trials using routine electronic health records: putting them to the test. *Br. Med. J.*, **344**, article e55.
Staniszewska, S., Denegri, S., Matthews, R. and Minogue, V. (2018) Reviewing progress in public involvement in NIHR research: developing and implementing a new vision for the future. *BMJ Open*, **8**, article e017124.
Stiller, C. A. and Draper, G. J. (1989) Treatment centre size, entry to trials, and survival in acute lymphoblastic leukaemia. *Arch. Dis. Chldhd*, **64**, 657–661.
Titmuss, R. M., Oakley, A. and Ashton, J. (1997) *The Gift Relationship: from Human Blood to Social Policy*. New York: New Press.
Vandenbroucke, J. P. (2009) The HRT controversy: observational studies and RCTs fall in line. *Lancet*, **373**, 1233–1235.
Whitty, C. J. M. (2017) Harveian Oration 2017: triumphs and challenges in a world shaped by medicine. *Clin. Med.*, **17**, 537–544.
Yates, F. (1963) Sir Ronald Aylmer Fisher, 1890-1962. *J. R. Statist. Soc.* A, **126**, 168–169.
Yuan, Y., Bachl, F. E., Lindgren, F., Borchers, D. L., Illian, J. B., Buckland, S. T., Rue, H. and Gerrodette, T. (2017) Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales. *Ann. Appl. Statist.*, **11**, 2270–2297.

Vote of thanks

David Spiegelhalter (*University of Cambridge*)
I thank Deborah for a very fine presentation. It is illuminating to see such continuity from these historical giants—Farr, Nightingale, Booth and Guy—to the present day: although we now have more data, more computing power and tackle different problems, the basic challenges stay much the same whether, in the useful division that was made by the President, it is to do with the public good, developing statistical science or developing statistical capacity.

The quotes from Florence Nightingale are, as usual, deeply impressive. She says we should start planting mustard seeds, thus pre-empting every self-help book that advises just to get going on projects. She was known for starting campaigns, giving them her best effort and, if they did not take off, then she would stop and use her energies elsewhere. Her other comment about pigeonholes says that conclusions 'must be tested by results'. I think that she would have been riveted by the network of What Works Centres that have been set up to review the evidence for policies in health, crime, education, homelessness and so on—the UK is world leading in this area. More than a quarter of state-funded schools have participated in randomized trials run by the Education Endowment Foundation (Major and Turner, 2017)—an extraordinary development that might have impressed the Victorians.

But often the available evidence is simply not very good, and I believe that the big challenge to statistical science in the future is the appropriate handling of observational data. What can we say when we have *all* the data, and not a survey with quantifiable sampling error, but we know that there are systematic biases, either reducing internal validity (lack of rigour) or external validity (lack of relevance)? When we know that our data have problems and yet we want to draw conclusions, is it enough just to list *caveats*? I do not think that this issue is ever really taught: in statistical courses we assume, say, that $X_1, \ldots, X_n$ are independently and identically distributed random variables from a normal distribution. But they never are! And such assumptions are likely to be increasingly less reasonable.

We know the problems of ignoring the biases in data: a topic that machine learning and artificial intelligence are wrestling with now. It is something that the statistical profession really must take seriously, both because official statistics will be increasingly based on administrative rather than survey data and because of the increased use of routine, 'real world', or 'found' data in healthcare evaluation. My particular interest is twofold: first, the development and appropriate use of 'quality-of-evidence' scales such as GRADE (Balshem *et al.*, 2011); second, the extent to which background knowledge and judgement can be used to quantify the potential biases. Lash *et al.* (2014) recently reviewed all such attempts, going beyond a list of rather vague *caveats* about limitations in the sample and so on, but instead trying to say how big those biases might actually be. I can point to work that we carried out with Rebecca Turner (Turner *et al.*, 2009) that provides a set of tools for quantifying potential biases as a distribution.

As statistics becomes more politicized, I believe that our profession must lead the way in demonstrating trustworthiness (Spiegelhalter, 2017), which is the first 'pillar' of the new code of practice for official statistics (UK Statistics Authority, 2018). As a recent example of untrustworthy use of statistics, recall the recent headlines claiming an 'insect apocalypse', numbers declining at 2.5% per year, no insects in a century, and so on (Guardian, 2019). This was based on a peer-reviewed meta-analysis (Sánchez-Bayo and Wyckhuys, 2019), in which the studies were almost all in Europe and yet global claims were made. But even more concerning were their search terms for the systematic review, which included 'insect', 'survey' and 'decline'! So they found all the papers about declining insect numbers and, hardly surprisingly, they concluded that insects are declining. Which they probably are, and I am sure that there is a problem, but this study is still not trustworthy.

We must resist the temptation towards what is known as 'white hat bias' (Wikipedia, 2019), which refers to the situation when someone feels they are on the 'right' side, and so they feel permitted to cut scientific corners and to overclaim. This is usually associated with exaggerated public health claims but also is seen in climate and environment. Even when we feel that we are one of the 'goodies', we must fight this tendency and call it out when we see it—we must be trustworthy; otherwise why should anyone take any notice of what we say?

On the topic of capacity, in this new era of data science, we clearly need a huge number of people who are capable of handling data. The Royal Statistical Society has done fine work in being open to data science, and I believe should be its professional body. But, as I pointed out above, traditional statistics education has generally been a (possibly watered-down) version of 'let $X_1, \ldots, X_n$ be independently and identically distributed…' etc., and I think that there needs to be some serious thought about what really needs to be taught to analysts rather than full statisticians.

I believe that the President, with her reflections on history, has demonstrated what a fine profession we belong to, which I recently referred to as demonstrating

> 'endearing traits of pedantry, generosity, integrity, and desire to use data in the best way possible'

(Spiegelhalter, 2019). It gives me huge pleasure to propose the vote of thanks.

Peter J. Diggle (*Lancaster University and Health Data Research UK, London*)
Professor Ashby's short history of the Society included many details of which I was previously unaware. In statistics as in other walks of life, we can learn from the past. One thing that strikes me here is a temptation to say that we are coming full circle, in the following sense. In the early years of the Society, data, and only data, were considered to be the Society's business. Clearly, the intention was to inform the discussion of important public policy issues, but there seems to have been an implicit assumption that the statistician's role was only to

> 'arrange, digest and publish facts… these facts being for the most part arranged in tabular forms and in accordance with the principles of the numerical method',

leaving it to subject matter experts to conduct what we might nowadays call scientific inference. Around 140 years since the granting of the Royal Charter, there is a tendency among some enthusiasts of 'big data' to think that context-free algorithms, if fed enough data, can become omniscient. For my part I very much hope that the near future of our discipline is closer to Judea Pearl's pithy riposte, to whit:

'You are smarter than your data. Data do not understand causes and effects, humans do'

(Pearl and Mackenzie, 2018).

What is undeniable is that the widespread use of phrases like 'big data', 'visualization' and 'data science', in harness with the work of first-rate communicators such as journalist and Honorary Fellow Tim Harford, Professor Ashby's predecessor Sir David Spiegelhalter and the late Hans Rosling have improved the public image of our subject to the extent that introducing yourself as a statistician at a social gathering no longer automatically ends the conversation. We must continue to capitalize on this at every opportunity... but especially when we are communicating with young people. Professor Ashby's references to our first woman Fellow, Florence Nightingale, are especially timely, as 2020 will be the bicentennial of her birth. At Health Data Research UK we have been giving some thought to the establishment of a 'Nightingale Lectureship Team' to prepare material on health data science and to deliver this material to primary and secondary schools across the UK. This would be a shameless steal of the idea behind the Guy Lectureship, but I would like nothing better than for it to become a joint venture between Health Data Research UK and the Society.

Another ambition for the near future is contained in an addition to tonight's set of quotes from history, in this case by the English writer H.G. Wells (1866–1946):

'Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write!';

but not, apparently, and despite the Society's best efforts, an essential attribute for a successful career in politics.

I have failed in my duty to criticize some aspect of Professor Ashby's address. This makes it especially easy for me formally to second the vote of thanks.

The vote of thanks was passed by acclamation.

References

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., Norris, S. and Guyatt, G. H. (2011) GRADE guidelines: 3, Rating the quality of evidence. *J. Clin. Epidem.*, **64**, 401–406.

Guardian (2019) Plummeting insect numbers "threaten collapse of nature". *Guardian*. (Available from `https://www.theguardian.com/environment/2019/feb/10/plummeting-insect-numbers-threaten-collapse-of-nature`.)

Lash, T. L., Fox, M. P., MacLehose, R. F., Maldonado, G., McCandless, L. C. and Greenland, S. (2014) Good practices for quantitative bias analysis. *Int. J. Epidem.*, **43**, 1969–1985.

Major, L. E. and Turner, J. (2017) Schools have learnt what works; now it's time to do what works. *News*. Education Endowment Foundation, London. (Available from `https://educationendowmentfoundation.org.uk/news/schools-have-learnt-what-works-now-its-time-to-do-what-works`.)

Pearl, J. and Mackenzie, D. (2018) *The Book of Why: the New Science of Cause and Effect*. London: Allen Lane.

Sánchez-Bayo, F. and Wyckhuys, K. A. G. (2019) Worldwide decline of the entomofauna: a review of its drivers. *Biol. Conservn*, **232**, 8–27.

Spiegelhalter, D. (2017) Trust in numbers. *J. R. Statist. Soc.* A, **180**, 949–965.

Spiegelhalter, D. (2019) *The Art of Statistics: Learning from Data*. London: Penguin UK.

Turner, R. M., Spiegelhalter, D. J., Smith, G. C. S. and Thompson, S. G. (2009) Bias modelling in evidence synthesis. *J. R. Statist. Soc.* A, **172**, 21–47.

UK Statistics Authority (2018) Code of practice for statistics. UK Statistics Authority, London. (Available from `https://www.statisticsauthority.gov.uk/code-of-practice/the-code/`.)

Wikipedia (2019) White hat bias. In *Wikipedia*. (Available from `https://en.wikipedia.org/w/index.php?title=White_hat_bias&oldid=901314827`.)