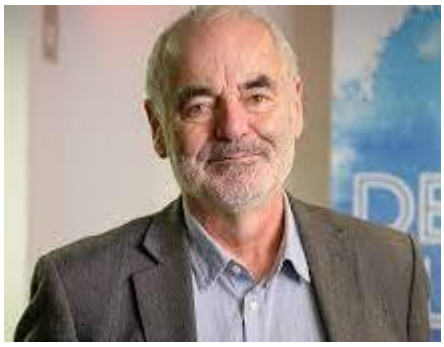
FIFTY-SEVENTH CONVOCATION ADDRESS

Professor Sir David J Spiegelhalter FRS OBE
Chair, Winton Centre for Risk and Evidence Communication
University of Cambridge, UK

DEAN

1/31/23

INDIAN STATISTICAL INSTITUTE



Professor Sir David Spiegelhalter FRS OBE is Chair of the Winton Centre for Risk and Evidence Communication in the Centre for Mathematical Sciences at the University of Cambridge, which aims to improve the way that statistical evidence is used by health professionals, patients, lawyers and judges, media and policy-makers. He was elected as President of the *Royal Statistical Society* and took up the position on 1 January 2017. He is among the many ISI Highly Cited Researchers. Google Scholar's current list of top-cited statisticians places him in the 34th position globally and in the 12th position in the list of biostatisticians. He was knighted in 2014 for services to medical statistics, and became a Non-Executive Director of the UK Statistics Authority in 2020.

Career

He started his career as a research assistant in Brunel University in 1976, followed by a visiting lectureship at the University of California, Berkeley during 1977–78. His doctoral thesis (awarded in 1978 by the University of London) with Adrian F M Smith as supervisor, was on *Adaptive inference using finite mixture models.* After his PhD, during 1978-81, he was a Royal College of Physicians Research Assistant at the Department of Mathematics, University of Nottingham.

From 1981 to 2012, he was at the Medical Research Council Biostatistics Unit at Cambridge, initially as a Research Assistant and later on as a Statistician. Between 2007 and 2012, he divided his work between the Cambridge Statistical Laboratory and the Medical Research Council Biostatistics Unit.

He has been an honorary lecturer at the University of Hong Kong since 1991. He has also been a consultant for GlaxoSmithKline, Novartis and the World Anti-Doping Agency. He played a leading role in the public inquiries into children's heart surgery at the Bristol Royal Infirmary, the murders by Harold Shipman, the independent inquiry on Mid Staffs Trust and, more recently, the infected blood inquiry.

From October 2016, he has been Chair of the Winton Centre for Risk and Evidence Communication. This Centre, generously funded by the David and Claudia Harding Foundation, is based in the Department of Pure Mathematics and Mathematical Statistics in Cambridge, and is dedicated to improving the way that quantitative evidence is used in society.

From October 2007 until December 2018, he was the Winton Professor for the Public Understanding of Risk in the Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge. Working with a small team, he was responsible for creating a website titled *Understanding Uncertainty*. This is an educational resource featuring the use of probability and statistics in everyday life, and makes extensive use of animations to help 'tell the story' of the data.

Research

His research interests in Statistics include

- Bayesian approach to clinical trials, expert systems and complex modelling and epidemiology
- Graphical models of conditional independence. He wrote several papers in the 1980s that showed how probability could be incorporated into expert systems, a problem that seemed intractable at the time. He showed that while frequentist probability did not lend itself to expert systems, Bayesian probability most certainly did.
- Statistical software. In the 1990s He led the Medical Research Council team that developed WinBUGS (Bayesian analysis Using Gibbs Sampling), a statistical-modelling system allowing hierarchical prior distributions. WinBUGS and its successor OpenBUGS specifies graphical models using acyclic directed graphs whose nodes are random variables, which are updated using Gibbs sampling (an updating method for Markov chain Monte Carlo (MCMC) simulation). Earlier Bayesian software had required that the probability distribution for the observed data be an exponential family and that the prior be its conjugate distribution. Allowing flexible choices of prior distributions simplified hierarchical modelling and helped to promote multilevel models, which became widely used in epidemiology and education.
- General issues in clinical trials, including cluster randomisation, meta-analysis and ethical monitoring.
- Monitoring and comparing clinical and public-health outcomes and their associated publication as performance indicators.
- Public understanding of risk, including promoting concepts such as the micromort (a one in a million chance of death) and microlife (a 30-minute reduction of life expectancy).
- Media reporting of statistics, risk and probability and the wider conception of uncertainty as going beyond what is measured to model uncertainty, the unknown and the unmeasurable.

Publications

He has more than 300 publications in peer-reviewed journals, apart from numerous book chapters, articles in proceedings and non-refereed articles. He has authored/co-authored several books. His bestselling book, *The Art of Statistics*, was published in March 2019, and *Covid by Numbers* came out in September 2021.

Media Presence

He has considerable media presence, and he has utilized it to enhance public perception and understanding of probability and statistics in everyday life, as well as the notions of risk and uncertainty. He has contributed a number of articles to newspapers, magazines, online sites and blogs. He has appeared in numerous shows on radio and television.

In March 2020, he launched a podcast called *Risky Talk* where he interviews experts in risk and evidence communication on topics like genetics, nutrition, climate change and immigration. He presented the BBC4 documentaries *Tails You Win: The Science of Chance*, which described the application of probability in everyday life, as well as the award-winning *Climate Change by Numbers*. He appeared on *Desert Island Discs* in 2022, and in 2011, he was placed 7th in an episode of BBC1's *Winter Wipeout*.

His home page is http://www.statslab.cam.ac.uk/~david/ and he is on Twitter as @d_spiegel.

Honours and Awards

The list includes

- Guy Medal in Gold, Royal Statistical Society (2020)
- Michael Faraday Prize of the Royal Society (2020)
- Knighthood for services to statistics (2014)
- Weldon Memorial Prize and Medal, University of Oxford (2010)
- Honorary Professor of Biostatistics at University of Cambridge (2006)
- Elected Fellow of the Institute of Mathematics and its Applications (2006)
- Officer of the Order of the British Empire (OBE) (2006)
- Fellow of the Royal Society (2005)
- Guy Medal in Silver, Royal Statistical Society (1994)
- Award for Outstanding Statistical Application, American Statistical Association (1990)
- Guy Medal in Bronze, Royal Statistical Society (1985)

in addition to Honorary Doctorates from several universities including Plymouth University, Heriot-Watt University and Aalborg University (Denmark).

Indian Statistical Institute

FIFTY-SEVENTH CONVOCATION ADDRESS

Professor Sir David J Spiegelhalter FRS OBE
Chair, Winton Centre for Risk and Evidence Communication
University of Cambridge, UK

First of all, I would like to say how delighted I was to receive from Professor Bandyopadhyay the kind invitation to be the Guest of Honour for the ISI Convocation, and to Amita Pal, Dean of Studies, for facilitating my trip. It really is a great honour to be here, and I feel humbled when I look at the previous people who have had this role. Thank you so much.

It is also wonderful to be back here in Kolkata. This is my fourth visit. My first was in 1984 when they were building the Metro. I admit that I am obsessed with public transport, and Kolkata is therefore a dream destination: the last time we were here we treated ourselves to a guided tour of the huge range of options - riding on the trams, buses, ferries, Metro, autos, everything. Complete heaven for someone like me.

This is also a great opportunity to come to India again. Before the pandemic, my wife and I came nearly every year for a few weeks travelling. We've been to so many fine places: Amritsar, Jaipur, Jaisalmer, Dharamsala, Goa, Hampi, Mysore, but have only just touched the surface. Statisticians know that averages are deeply inadequate descriptions – we know we need to understand variability. And India has more than its fair share of variability!

It is my first visit to the ISI, even though of course I have known about it, and its illustrious history, since I started doing statistics and learned about Mahalanobis. I do in fact have a few things in common with The Professor: we have both been awarded an OBE, FRS, and the Weldon Prize, and enjoy punting in Cambridge. But sadly I cannot cover the huge range of his skills and interests, which lay at his remarkable vision of ISI as a centre cross-disciplinary collaboration. Which means I must apologise in advance that my talk is going to focus on statistics, as I know the ISI covers so much more.

Maybe I should give some details of my own journey with statistics. I studied mathematics at Oxford University and really enjoyed the pure maths, but I have to admit that by the end of the second year it all got too difficult - maybe many of us have a ceiling of abstraction on which we eventually bang our heads.

But I was extremely fortunate in that my tutor at Oxford was a young statistician called Adrian Smith, now professor Sir Adrian Smith, the President of the Royal Society. He aroused my interest in probability and statistics, often through heated discussions in the pub. He was a convinced subjectivist Bayesian, and was in the process of translating into English de Finetti's book *The Theory of Probability*, which starts with the powerful statement - *Probability does not exist*. This is not the time nor the place to get into arguments about what probability really is. But the passion that I felt as a 20-year old for the basic ideas behind probability and

statistics has never really left me, and I am happy anytime to join in and argue with anyone who claims probabilities objectively exist, except perhaps at the quantum level.

I then did methodological research, largely in Bayesian methods in biostatistics, and was fortunate to be involved in the revolution in Bayesian methods around 1990 when simulation methods, specifically Markov chain Monte Carlo, suddenly allowed Bayesian analysis to be feasible in realistically complex models. Since then of course Bayesian methods have become very widely used in a huge range of applications.

When I started teaching undergraduate mathematicians at Cambridge University, in the same Tripos that Mahalanobis began before he switched to Physics, I gave a fairly traditional course rather similar to the ISI undergraduate syllabus. I do think that there needs to be a core of mathematically-trained statisticians that can take forward new methodology.

While I was working at the MRC Biostatistics Unit I started becoming involved in high profile issues, in particular two public inquiries into scandals in the UK. The first concerned the mortality rate of babies undergoing heart surgery at Bristol Royal Infirmary, which had been reported by a whistle-blower and then confirmed by a review of the data. I headed a team of statisticians who estimated there had been 35 more deaths than would have been expected in an average hospital at that time. This required giving evidence to the inquiry with the families of the victims sitting in the front row, which required great care and thought about how to discuss the statistics of mortality rates without sounding callous and harsh.

Then a few years later I was in the statistical team working on the inquiry into Harold Shipman, who was a general practitioner who over a 20-year period murdered at least 250 of his patients and possibly up to 400. He had finally been caught and imprisoned, but we were asked whether he could have been caught earlier. By adapting industrial quality control methods, specifically the sequential probability ratio test, we concluded that if somebody had been looking at the data an alarm could have been given about his excess death rate after only about 40 deaths which should have caught him and saved hundreds of lives. But nobody was looking at the data. This taught me that asking the right questions and getting the right data could be more important than the exact methods that are used for the analysis.

Again we gave evidence with the victims' families in the front row. This was yet again repeated recently at the UK Inquiry into the thousands of people who were infected with HIV and Hepatitis C through being given infected blood in the 1970s and 1980s.

I started talking about these issues in public, and found I enjoyed demonstrating the power of statistics in the real world. I was then fortunate enough to get a new job as (the one and only) Professor for the Public Understanding of Risk, which essentially meant I became a 'performing statistician'. I gave huge numbers of talks to schools and other audiences, wrote 'popular' books, (one of which, *The Art of Statistics*, actually became popular!), fronted BBC TV documentaries, made radio programmes, gave interviews, wrote articles in papers and magazines, and got to know journalists and the way the media works.

And then COVID came along. This meant constant demands from the media for explanation of the huge number of statistics with which the public were being bombarded. This was very

challenging. First, there are many complex statistical issues, for example the definition of a COVID death, testing regimes, false positives and negatives, excess deaths, and so on. A colleague and I had a weekly column in one of the main Sunday newspapers through the whole of 2021, each covering a difficult statistical question in 350 words. So here is a challenge. Can you explain to a non-technical audience why, from June 2021, most people in the UK who died with COVID have been fully vaccinated, and why this was not a sign that the vaccines were useless, or worse?

There has been a lot of misinformation about these kinds of statistics, with President Bolsonaro of Brazil using UK data to claim that the vaccines were giving people COVID. But some statistical common sense, or Bayes theorem, tells us that when nearly everybody has been vaccinated, and the vaccine does not provide perfect protection, then the 'breakthrough' deaths can outnumber those in the unvaccinated group, even though the risk is higher if you haven't been vaccinated. A good analogy helps. Most people who die in car accidents are wearing seat belts, but it doesn't mean that seat belts are not an effective protection.

The pandemic demonstrated to everyone how important statistics are, both for those making policy and for the public understanding of what is going on. But we know that statistics can be used, or misused, to support extreme claims, especially on social media, and it can be challenging to keep a middle way between polarised opinions. There is often pressure from the media to take a strong view, which I think should be resisted by statisticians and other scientists. Also, should one engage with extreme sceptics, either those who question vaccines or, on the other side, COVID itself? Personally, I don't bother. People with extreme views tend to have little regard for rules of engagement, and will always come up with yet another claim which would require a lot of work to show it was wrong.

The pandemic also convinced me of the value of trustworthy communication. This requires trying to *inform* people rather than just trying to *persuade* them, giving them both the pros and cons, the winners and losers, of any policy, and acknowledging the inevitable uncertainties. Our team was asked to produce an infographic about the AstraZeneca vaccine, and we showed both its benefits in reducing severe COVID and its potential harms in terms of blood clots. The graphic clearly demonstrated that for people my age the benefits hugely outweighed the risks, but for younger people they were more finely balanced. So when it was announced that younger people would not be getting that vaccine, it was well accepted by both the media and the public

It was not my role to criticise the government, although on rare occasions I did speak out about the use of statistics in public briefings, when there were lots of big numbers being thrown around about tests, infections, and deaths, without appropriate context or interpretation. I called this *number theatre*, which I think is the wrong kind of statistical performance

I am also a non-executive director of the UK Statistics Authority, which supervises the Office for National Statistics. We were prescient in the UK to set up a proper population-based COVID infection survey early in the pandemic, which has enabled us to monitor infections, antibodies, symptoms and many other aspects of the pandemic. It's notable that this is based

on a close collaboration between the Office for National Statistics and Oxford University, and that the most advanced academic statistical methodology - multi-level regression and post stratification (MRP) - is being used in its analysis. Fully Bayesian, I should say.

I used to think that official statistics and surveys were not a very interesting part of the subject but have now completely changed my mind. They are both important and challenging, requires deep methodology combined with a real interest in influencing policy and raising public understanding. This is exactly what Professor Mahalanobis demonstrated so many years ago, in his vision of surveys being vital for planning policy. As Sir Ronald Fisher said on the occasion of the first convocation of the Indian Statistical Institute: 'I need hardly say that I refer to the emergence of a statistically competent technique of sample survey, with which I believe Professor Mahalanobis' name will always be associated.'

It seems to me that India still provides the ideal platform for such interdisciplinary work, given its strong history in statistics, both official and methodological, its skills, and its digital familiarity. And of course, its large and varied population provides an ideal basis for biostatistical research.

I hope I have shown that statistics plays and increasingly important role in the modern world, in trying to establish what is actually true (whatever that means). But the subject must adapt. I see two specific challenges: the growth of data science and machine learning and raising the profile of statistics in public life.

First, I believe that that statisticians should welcome the rise of data science and machine learning, although it can feel strange to hear logistic regression being quoted as a machine-learning method. Some US statistics departments are renaming themselves as 'statistics and data science', and the Royal Statistical Society is founding a new data science journal. I do think it is important that statistics training includes exciting ideas arising from the analysis of really big data sets. Even my university, which can sometimes feel medieval in its approach, has had to change and develop courses on modern methods in statistics. I note the ISI winter school in deep learning is going on even as I speak.

Second, if statistics is going to raise its profile, it requires statisticians to stick their heads above the parapet and engage. This does not mean taking sides in political or social arguments – it means trying to raise the level of the argument by making clear the strengths and limitations of available evidence.

Old people are supposed to give wise advice. I fulfil one of those criteria, in that I am old, but I doubt my wisdom. But here goes anyway.

If any of you feel enthusiastic about communication, I would recommend the following. First, start by making academic presentations as interesting, and even entertaining, as possible. This can mean using images whenever possible, and even humour if it suits you - stats can be funny! Crucially, stick to the time allowed, pay attention to the audience, and get feedback from a critical friend.

Second, try explaining your work, or statistics in the real world, to your friends or family – I always imagine I am talking to my old aunt, who is no longer with us, and to be honest was not very good at understanding things. This forces you out of using jargon, as you need to describe complex ideas without using any of the words with which we are so familiar. We have trained statistical communicators by forcing them to explain ideas like p-values in tweets, and of course they are not allowed to use phrases such as statistical significance or confidence intervals.

Third, start to notice good journalism, and how good communicators can explain complex topics in simple language. I was lucky to get good media training and so learn how to deal with being interviewed on radio and television, but I still get nervous and still make mistakes. I used to think I had to answer every question, but learned during the pandemic that it was fine to say that something was not my job and they'd have to ask somebody else - in fact at times I was begging the interviewer to ask me something I actually knew about. Audiences seem to respect that.

Fourth, I have to admit that it helps to be reasonably successful in your field, and so have both the experience and status to present oneself as a suitable 'expert'.

Finally, when communicating to the media and the public, maybe the biggest lesson is that things will go wrong. I have forgotten to unmute myself when appearing live on prime-time TV. Some resilience is needed, particularly as some people love to criticise. But, and of course this is a general lesson in all endeavours, learn the lessons, and try again.

To conclude, it is wonderful to see you all having worked so hard to gain your qualifications and ready to embark on the next stage in your career. On the whole, mathematicians, statisticians, engineers and scientists just want to do a good job and help others to understand the world. These are honourable professions, and I hope you are as proud as I am to be part of them. And of course being an alumni of the ISI means that you have the opportunity to reflect its fine heritage and take forward its vision into the world.

I wish you every good fortune in what you choose to do.
