

A conversation with David Colquhoun, London

# “Bibliometricians are the Curse of the Age”

Once a successful pharmacologist, working on ion channels with amongst others Nobelist Bert Sakmann, David Colquhoun now focusses on calling out shortcomings in the scientific world. Unfair bibliometrics, academic bullying, pseudoscience or the reproducibility crisis are just a few issues he worries about but “It’s in our hands to fix it”.

**L**ab Times: What made you decide to start your prize-winning blogsite “DC’s Improbable Science” and to actively write about wider issues affecting science?

**Colquhoun:** The thing that got me into blogging in 2002 was when Imperial College London tried to take over University College London. We collected signatures from various prominent people but what turned out to be more important was tell-

chancellor told them, “I know I said there won’t be any redundancies, but of course there will, but don’t worry they will not be in Imperial College.” And somebody at the meeting wrote this up, distributed it around his department and within five minutes, I had two separate copies, and ten minutes later it was on the web and public knowledge. The next day the whole thing folded. They were such idiots.

But that got me into blogging because, suddenly, I realised that you could sit there in front of the computer and hit a key and actually affect things that were happening in the real world. And this was still rather novel. So, I tried it again with politics and quackery, and other such things.

*In the Guardian, you wrote an article, in which you described the past 30 years “as an Age of En-darkenment,.. a period in which truth ceased to matter very much, and dogma and irrationality became once more respect-*

*able” (15/08/07). You’ve written critically about alternative medicine and the role UK universities and government have played in promoting it.*

**Colquhoun:** Yes, I was commenting about homoeopathy, which is obviously such a low-hanging fruit because the homoeopathic pills contain nothing. Therefore, a trial that gives a positive result must be a false positive. But there are far more dangerous consequences than a few batty homoeopaths believing things that you wish to

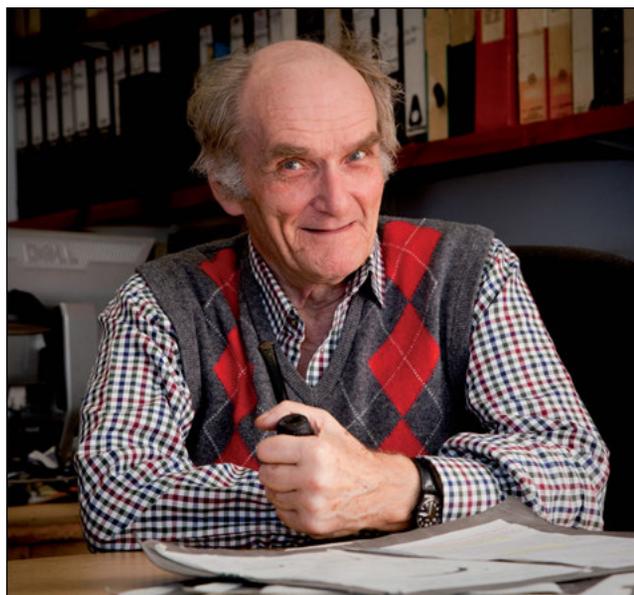
be true that aren’t. Most of my talks begin with a slide, which shows the UCL Quad on March 20<sup>th</sup> 2003 at the start of the second great march to stop the war in Iraq. Each time I talk about this, I say there are worst consequences of believing things that aren’t true than homoeopathy (“when people delude themselves into believing that we could be endangered at 45 minutes’ notice by weapons of mass destruction”). It seems to spread to all reaches of life.

*It seems quite incredible that BSc degrees are being taught in UK universities for subjects like homoeopathy and aromatherapy, subjects that have no scientific basis.*

**Colquhoun:** I had a commentary in *Nature* about that in 2007 (446, 373–4). I was as astonished as anybody to realise you could get a BSc in homoeopathy. What’s going on? I think the first post that had some impact was one about amethysts emitting yin energy (*DC’s Improbable Science*, 23/04/08). Some crystal therapist taught this to first year University of Westminster students, some of whom were sufficiently incensed that they contacted me and I posted the lecture slides. The head of Westminster at the time was supposed to have been a geomorphologist. So I wrote to him and asked him what his opinion was, as a geologist, that amethysts emit high yin energy but, of course, he didn’t reply. They just ignore this. It is incredibly rude.

*This brings us to other problems in universities in Britain, for example, the whole question of bibliometrics that have distorted the way we assess and measure science. With the rise of ‘managerialism’ a lot depends upon what are, in effect, ‘false statistics’.*

**Colquhoun:** Yes, it comes back to the statistics. Bibliometricians are really the curse of the age. All they do is to correlate



ing people what was happening. There were meetings between the departments at UCL and at Imperial College to discuss ways of merging their efforts. We were told that these initial meetings would report to two committees then everything would be published openly and transparently, but everybody knows what happens when things have been through two committees. Instead, what happened was that people, who attended the committees, would send me their raw unprocessed minutes and I would stick them on the web. After five weeks, the whole thing fell through because at a meeting of Imperial College Council, their vice

**“Suddenly, I realised that you could sit there in front of the computer and hit a key and affect things in the real world.”**

one silly metric with another. Sometimes they find a correlation and sometimes they don't. Regardless of whether they find a correlation, none of them really answers the question 'What encourages good science?'. I wrote an article that presents a different way to look at the problem, which is to take a scientist who is universally respected in the field – I took Bert Sakmann as an example – and look at his publication record (*Physiology News*, Winter 2007; *DC's Improbable Science*, 3/8/07). I discovered that if you take the ten years, in which he was coming to fame, i.e. 1976 to 1985 – from the date of the first single channel paper up to our big paper together – he would have failed Imperial College's publication metric in six of those 10 years. In two of those years, he had no publications whatsoever. So, Imperial College might well have fired him. In 1991, he won the Nobel Prize [with Erwin Neher for their discoveries concerning the function of single ion channels in cells]. Fred Sanger would almost certainly have been fired by Imperial College or many other universities, these days [he was awarded two Nobel Prizes in Chemistry in 1958 and 1980].

Examples like that seem to me quite sufficient to show that trying to measure the quality of research by counting citations is nonsensical and will probably result in the firing of the best people. It seems to be based upon the premise that if you adopt harsh enough criteria, you can get a whole department full of Nobel Prize winners. But the fact is that you cannot. There are not enough of them to go around. Of course, winning the prize or doing some really important work is almost as much a matter of luck as talent. Most of the time, most people are not going to be wildly successful. They are going to do good work but they are not going to win big prizes. But you can't force people to become geniuses by saying we will fire you if you do not bring in £200,000 a year in grant money. Which is, of course, what Imperial College did most famously to Stefan Grimm, who committed suicide as a result (*DC's Improbable Science*, 1/12/14).

You wrote in 2007 about Imperial College's excessive demands of their scientists – high publication rates, good bibliometrics, big grant money. You pointed out that it is in

effect "bullying" of academics. It was seven years later, however, that Stefan Grimm committed suicide in response to just such pressures. Imperial College has simply continued?

**Colquhoun:** Yes, they took not the slightest bit of notice. In fact one of the bullies got a knighthood in 2012. It is bullying and it is also a great incentive for people to be dishonest and to take shortcuts. It is actually corrupting science. It doesn't result in many deaths but it certainly makes many lives very miserable. Stefan Grimm was 51 [and Professor of Toxicology in the Faculty of Medicine at Imperial College London]. He had stacks of publications, he might not have been Nobel Prize quality – I do not know the field well enough to say – but he had had some difficulty in getting grant money. Who the hell doesn't these days? He seemed to be doing perfectly well to most people. But these things are all ephemeral anyway – you can go from being second-rate to being a first-rater overnight or the other direction. It's a matter of luck largely, it's stochastic. You cannot punish people for it. These measures are so crude. It is statistically illiterate apart from anything else.

University rankings also distort things. In 1996, David Spiegelhalter wrote a paper about the uncertainty in rankings, including university rankings, and he showed there is no way you can tell the difference between the top 10 universities (*JR Stat Soc*, 159(3): 385-443). Yet vice chancellors kill people, to go up one place in the world university rankings. If they simply ignored these things, people would stop producing them. OK, they're a money-spinner for Times Higher Education, but if they were simply ignored, they'd go away. I do wish that would happen.

We both wrote about the scandal at Queen Mary University of London (QMUL) in 2012 (*DC's Improbable Science*, 29/6/12; *LT* 4/2012 p.20-25; *LT* online 4/07/12). In order to go up some places in the university rankings, QMUL demanded that its academic staff meet performance targets for levels of research funding, numbers and impact factors of papers. Those who did not meet these targets were sacked in order to bring in people who looked better on paper. You said this seemed like "scientific suicide" on their part

– how can they expect to get good scientists when they know that they, too, will be kicked out as soon as the university decides that they are not performing well enough?

**Colquhoun:** Well exactly. And who the hell would want to work at Imperial College London knowing that they may be kicked out in their mid-50s because they're not getting enough grant money? It seems really silly and also counter-productive for the university in the long run. Who the hell would want work under a regime like that? They kill their employees. Literally in one case.

I heard from Stefan Grimm's mother this summer. The Times Higher Education hadn't published his original e-mail, so I published it on my blog. Someone logged on to it every second from all over the world and the server went down for a few hours. It's now been viewed 200,000 times. But I didn't make any attempt to contact his parents. I didn't know how and it would have seemed intrusive. Then in July, I had a handwritten letter from his mother who is 80 and living in Munich. It was so moving. She was thanking me, because she said "most of what I learned about my son's death was from your blog". What was going on? The University sent a couple of short notes, the usual token messages that she didn't want me to publish. She seemed very grateful about my efforts and sent me a lot of his early drawings which I posted as a memorial on 25<sup>th</sup> September 2015 – the first anniversary of his death. I would not like people to forget that. But the Times Higher and the *Guardian* were not interested in it – they said it was old news. That's the way it goes with journalism. Unless it's topical, they don't care. But I think that people should not forget these things.

Presumably you had tenure when you became a lecturer?

**Colquhoun:** Well it was never a very formal thing but it was understood. Yes, I think it was Margaret Thatcher who formally removed tenure. Until recently, the practice was not much affected by that. However, if this current government reduces the funding for science, there is going to be carnage.

They'll be getting rid of a lot of lecturers and researchers, who otherwise had stable positions?

**Colquhoun:** I think so, yes. They may have no option. Without the cash, they can't afford to keep them. A lot is going to depend upon the spending review which is set for 25<sup>th</sup> November [the government expects £20 billion in departmental budget cuts for

the coming five years]. There could be out-of-work scientists all over the streets.

*Much university research is performed by people on short-term contracts. They hope one day to gain stable employment but you're saying this may be lost, too, e.g. by replacing lecturers with short-term staff? In one of your blogs, you spoke about the University of Warwick, where they get teaching done by people who are on the equivalent of "zero hours" contracts (DC's Improbable Science, 8/04/15).*

**Colquhoun:** Yes, that is awful. For a start, in my area, it's going to be hard to find any part-time persons, who understand the subject well enough to give lectures on topics like the binding-gating problem. Half of my colleagues don't understand it, never mind teachers who come in for short-term teaching. The quality is bound to suffer. And as you say, these people have no rights. It is like the casual dock labour that existed when I was a child in Birkenhead. People would turn up every day to see if there was a job. You cannot make a career that way. And you cannot expect people to be very committed either.

*You also wrote about the increasing lack of reproducibility of published research re-*

*sults. You mentioned two reasons for this. On the one hand, the well-known pressure issue but you also get back to statistics and how researchers use them to say that what they have found is true because it is "statistically significant".*

**Colquhoun:** 'P equals 0.045, therefore I have made a great discovery.' Yes that's an interesting process. Because although I have been interested in statistics for a long time, I have very rarely done tests of significance. So it had really escaped me. Also, I had been put off by the perpetual internecine rows among statisticians, between Bayesians and Frequentists. And I had rather dismissed the idea of interpretation of P values properly as being sort of Bayesian.

But then recently, screening tests have come into prominence and a lot of my friends have been very active in pointing out that some of these screening programs may do more harm than good because they produce so many false positives (DC's Improbable Science, 10/03/14). If you have a test that has a sensitivity of 95% and a specificity of 80% and you try to find a condition that is present in 1% of the population, then you get 86% false pos-

itives. That is a disaster. For a start, it will cost a lot of money and the false positives may, for example, have their breasts cut off unnecessarily.

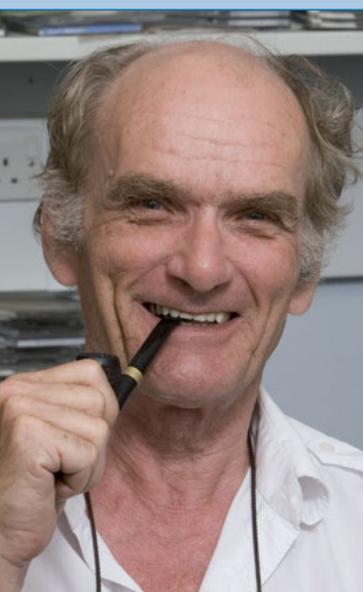
So, it occurred to me one day that this is analogous to the argument that you can apply to significance tests. This had not explicitly occurred to me before. I'm ashamed to say this as someone with such a long-standing interest in these things. The P value does exactly what it says on the tin, but what it says on the tin is not what you want to know. What you want to know is "if I claim to have made a discovery on the basis of the P-value, how often will I be wrong?" And it turns out that if the P value is a marginal one, 0.047 for example, and if you

*"The P value does exactly what it says on the tin, but what it says on the tin is not what you want to know."*

say I've made a discovery every time P equals 0.047, then you are going to be wrong at least 30% of the time. And much more than 30% if the hypothesis is an implausible one to begin with.

And I thought – why wasn't I taught this?

So, I put it on the blog (DC's Improbable Science, 24/03/14) and then I wrote a paper and put that on ArXiv. And during this time, I collected a lot of feedback and had a lot of discussions. In the end, I found a jour-



**David Colquhoun** (b. 1936) started work as an apprentice pharmacist in Birkenhead, Merseyside, going on to study pharmacy, with a specialisation in pharmacology at the University of Leeds. During his BSc, he developed an interest in statistics and random processes, publishing his first paper on 'Logic and the interpretation of observation'. Following his PhD at the University of Edinburgh, where he studied the binding of immunoglobulins to lung tissue, he became a lecturer at University College London (UCL) in 1964. His research did not go well at that time but he wrote a textbook on statistics – 'Lectures on Biostatistics' (now freely available at: [www.dcscience.net](http://www.dcscience.net)) and then began to work with Alan Hawkes,

a statistician at UCL, on "why the time constant for dissociation of a molecule from a receptor would be the mean lifetime of a drug-receptor complex". He says if he hadn't met Alan Hawkes, "my career would have been quite different" because in explaining this paradox, he became interested in single molecule behaviour.

He spent most of the 1970s at Yale University and the University of Southampton, returning to UCL in 1979. During that time, he extended Bernard Katz's invention of noise analysis to predict the spectrum for any arbitrary mechanism (*Proc R Soc Lond B*, 199,

231–62). In 1976, a groundbreaking paper by Erwin Neher and Bert Sakmann had announced the patch clamp technique for measuring single-ion channels. Prior to the patch clamp method, ionic currents were recorded as whole cell currents and only the average behaviour of a large number of channels could be observed. At the time, there was no theoretical framework for the interpretation of single molecule measurements, so Hawkes and he had to develop it from scratch (*Philos Trans R Soc Lond B Biol Sci*, 300, 1–59). David Colquhoun began a productive experimental collaboration with Bert Sakmann to apply the theory to single channel data and to formulate a likely quantitative model describing how the channel functions (*Nature*, 294, 464–6; *J Physiol*, 369, 501–57). Subsequently, Hawkes found an exact solution to the problem posed by the fact that many events are too short to be resolved (*Philos Trans R Soc Lond B*, 337, 383–404). This allowed a maximum likelihood fitting programme, HJCFIT, to be developed (<http://tinyurl.com/oxjc42s>). In 1985, he became Professor of Pharmacology at UCL and was elected as a Fellow of the Royal Society (FRS). He won the Humboldt Prize in 1990. Upon his retirement in 2004, he was made an Honorary Fellow of UCL and continues to publish research.

In 2002, he started an internet page which eventually became his blog 'DC's Improbable Science', in which he has written critically about many issues affecting science and universities, including alternative medicine, confused government thinking about science, the nonsense of metrics, and the abuse and misunderstanding of statistics in scientific research. In 2012, his blog site was awarded the UK Science Blog Prize by the Good Thinking Society. It is archived for preservation by the British Library.

nal to publish it – it took four goes – I think they either did not like the message or said that it was not original enough. I mean, it's a review, not original research (*R Soc Open Sci*, 1, 140216). It must have come along at the right moment because it has now had 11,000 pdf downloads and 87,000 full text views. It's had wider interest than anything else I've ever written. But it's really rather simple when you think about it. So I am talking about that quite a lot now. I was rather dreading talking about it in the UCL statistics department – they are obviously professional statisticians – but in fact there was no serious dispute. I still do not understand why elementary statistics courses do not teach it. Because that fact alone will probably account for quite a lot of the crisis of reproducibility. There are a lot of other reasons, too, but that seems to me an important one.

An example: I blogged about claimed transcranial magnetic stimulation will improve your memory (*DC's Improbable Science*, 2/11/14). There was a tweet from *Science* that was re-tweeted many times. Anything to do with memory and the brain tends to get lots of re-tweets. So I looked at the actual paper (*Science*, 345:1054-7). But it wasn't a paper about memory at all! It was a huge fMRI study. In one of the figures (4B), there was this little memory test. It was very crude with only three-time points. And the difference between them looked very unconvincing to me, but it came out as  $P = 0.043$ . This was the basis for *Science* tweeting this great discovery! It was utterly unreliable, in my opinion, and yet another reason why we do not need glamour journals. They are in competition with each other, so if *Science* thinks it can get something that *Nature* hasn't and promote it mercilessly, they think it will be good for their reputation. In fact, it will be bad for their reputation when there are papers like that. Unfortunately, that is not how the editors see it. They just see it as putting up their impact factors. So the journals also have a part to play in the corruption of science.

*Many researchers just seem to rely on their computer's statistics programmes to tell them when  $P < 0.05$ .*

**Colquhoun:** Of course, in the old days you had to do the statistics by hand. Now you just plug it into a computer programme, which you may or may not understand. It will crunch out a number and eventually you'll get  $P$  below 0.05 and publish. And

that is worrying. The trouble is, the public has realised this and this just adds support for the vaccine deniers, climate deniers, etc. They just say – well, half your stuff is wrong anyway. Nobody knows what fraction of climate science is wrong, or indeed physics. In my own particular narrow field of single ion channels, there have been differences of interpretation but I have never noticed any serious inconsistencies in the data. So, I do not think it is a big problem there. It has been mainly psychology, cancer studies and genome-wide association studies (GWAS), though at least GWAS has now started to correct for multiple comparisons.

The particularly bad area seems to be experimental psychology. I wrote my article before the Nosek study that showed only 36% of psychology studies are reproducible (*Science*, 349(6251). This is a disastrous figure for science! Awful, shameful. But what worries me is that I see lots of people defending it. That's baffling to me. How can anyone defend this? They say it's only the first step and other people will confirm or not confirm it later. Well, there's a huge waste of money to keep on doing the same study and in the end finding there was nothing to it. But, of course, if you made  $N$  big enough it would take you twice as long to publish it and you would be fired in the meantime. It really is deeply corrupting to the point where it is harming the image of science with the general public.

*In your recent blog on reproducibility, you said you do not think there is a need for training courses on research ethics for young scientists because the people who are really causing the problems are those higher up in the system – the senior scientists, the university vice chancellors, etc. (DC's Improbable Science, 14/04/15)*

**Colquhoun:** I think that's right, by and large. There has been a recent case where a junior postdoc was found to have cheated on some experiments – which he shouldn't have done – but I happen to know something about that case. The postdoc had been bullied by his lab boss into getting particular results. The postdoc should not have done it but I can understand, to an extent, why he did, because he was bullied into it. The boss didn't really understand the principles of the subject himself but he had his idea about the result that he wanted and

he was instructing the postdoc to find it. So, it was the boss who needed the ethical instructions.

*You have also noted the incredible competition for limited research money. Perhaps when faced with a difficult choice between keeping your job or, as you have said, losing your home, the honesty of your science may be*

*one of the first things to get left behind.*

**Colquhoun:** Yes, if you are told that unless you produce lots of papers or that you bring in £200,000 in grant money each year, you are going to be homeless, it is asking an awful lot of human nature for people to resist pressure like that. It is going to lead to dishonesty and, in the end, it will be bad for the universities. But not before a few more people have been made homeless or have killed themselves. Of course, not everybody at UCL, Imperial College or wherever is madly productive. What do you expect? There are only a handful of people who are. You can't make a department that has all of them in it.

*Unfortunately this message does not seem to have been understood by managers of universities, let alone politicians.*

**Colquhoun:** It is not understood at all. I know there is a limit to what you can spend on research but I think if it were reduced further now in the UK, it would cause carnage. We don't spend as much as some other countries as it is. But we'll see what happens in the spending review - I am nervous about it.

*Are you optimistic about how things are going, in general?*

**Colquhoun:** I'm optimistic to the extent that the web, in particular, has given voice to a lot of people who talk sense. People who didn't previously have a voice. Think what you could do in the 1990s. All you could do was write a letter to the *Times*. Now, there are people holding meetings, making videos, pointing out these things. To that extent, I think that things are better than they were. At least the problems have been recognised now, so I hope that will reverse. We'll see. It's in our own hands to fix it.

INTERVIEW: JEREMY GARWOOD

[Read the full interview on our website: www.labtimes.org](http://www.labtimes.org)