



Science and Technology Committee

Oral evidence: Reproducibility and Research Integrity, HC 606

Wednesday 15 December 2021

Ordered by the House of Commons to be published on 15 December 2021.

Watch the meeting

Members present: Greg Clark (Chair); Aaron Bell; Katherine Fletcher; Mark Logan; Rebecca Long Bailey; Graham Stringer.

Questions 72 - 177

Witnesses

[I](#): Dr Ben Goldacre, Senior clinical research fellow at the Centre for Evidence-Based Medicine, the University of Oxford; and Dr Jessica Butler, Analytical Lead and Research Fellow, University of Aberdeen.

[II](#): Dr Ritu Dhand, Chief Scientific Officer, Springer Nature; and Dr Elizabeth Moylan, Publisher, Research Integrity and Publishing Ethics, Wiley.

[III](#): Richard Horton, Editor in Chief, *The Lancet*; Viscount Ridley, Co-author, *Viral: The Search for the Origin of Covid-19*; and Dr Alina Chan, Co-author, *Viral: The Search for the Origin of Covid-19*.

Written evidence from witnesses:

– [Add names of witnesses and hyperlink to submissions]



Examination of witnesses

Witnesses: Dr Goldacre and Dr Butler.

Q72 Chair: The Committee is now in session. The Science and Technology Committee continues our inquiry into research reproducibility and integrity. We are very pleased this morning to welcome our first two witnesses, Dr Ben Goldacre, senior clinical research fellow at the Centre for Evidence-Based Medicine at the University of Oxford, and Dr Jessica Butler, analytical lead and research fellow at the University of Aberdeen. Thank you very much indeed to both of you for joining us this morning.

Perhaps I could start with a question to Dr Butler. Could you describe the role that research institutions, including universities, currently play in the publication of academic research, bearing in mind that later this morning we will hear from publishers? How does that differ from how they should play that role in an ideal world?

Dr Butler: It is a great question. Thank you for the invitation. Apologies, I lost my voice just in time for this testimony.

Chair: We can still hear you, so thank you for persevering.

Dr Butler: The vast majority of research that gets done in the UK gets done at university. Barring some commercial enterprise, primary research in the UK gets done by researchers in the universities. Due to historical reasons, the primary way we output our research back into the public sphere is through publications and academic journals—literally magazines. We spend a year and a half doing the research. When we finish it, we write it up in the form of a magazine article and send it off to be peer reviewed—that is, error checked—by these science magazines. That is the way it has worked for about 150 years.

The peer review is organised by those magazines, but, somewhat oddly, we, the researchers with the expertise, do not organise the error checking or the quality control of our own research; we outsource that to those magazines. That gets done after the research has finished. Rather than inspecting the quality of the car that you are building all along the process, you wait until the car is finished, you drive it out of the factory, you deliver it to the car salesman, and you ask them to check the quality of the research.

So you can see that there are a number of ways we could improve this that are obvious, easy steps. Especially in this modern day and age of digital output, I do not think we need to depend on these journals as much as we do.

Q73 Chair: Thank you, that is very clear. You mentioned that journals organise the process of peer review, of taking referees. Typically, those referees are academics employed in universities, are they not? Is the distinction quite as sharp as it might appear—that you have these separate magazines, as you call them?



Dr Butler: What is interesting is that peer review—quality control for all of our academic research—is done by volunteers. The journals literally send out emails saying, “Could you please read this article and tell me if it looks okay?” They send out dozens and hope for the best—that somebody volunteers their time to do this. Academics do the quality control, but I get millions of emails: “Could you please read this article and tell me how it looks?”

I get no reward in my job for doing this. My bosses don’t care if I spend time on this. In fact, they prefer I do other things—the research I am paid to do, or to go out and get grants. You are depending on volunteers who are unpaid.

The other interesting thing is that, if I get a bit of research to peer review that I think is no good—I find a devastating problem in the way they ran their experiments—and I write this up and send that peer review back to the journal, and they say, “Okay, excellent, we won’t publish this research,” the researchers are free to go back and submit that research to any other journal, and I won’t hear of it again and my negative review disappears. I don’t own it. The magazine owns that review. It never goes public.

It is very easy to find a less ethical journal that will put that research out there, and that time and effort that I put into the excellent negative review that I put forward disappears from the record. There is a lot to be said for us, the UK researchers, organising the quality control a lot further upstream and making it very public.

Q74 **Chair:** Thank you. Let me press you on that thought. If it were to be further upstream, should it be done by the universities as institutions, or should it be done by learned societies? Of course, some of the journals are the journals of learned societies. Do you have in mind a way that it could be better done upstream?

Dr Butler: I am keen to move that peer review—that quality evaluation—very far towards the beginning of the process. If I am writing a grant proposal today to the UKRI—to the Government—to fund my research, there is no way I could get funding to hire someone to critique my work; it is not a position that exists in academia. I think we should have that. Let us put that aside.

There is a method that currently exists for getting peer review earlier. You heard two weeks ago about registered reports. It is a name I hate. All it means is that I develop my research proposal. I do not collect any data; I just write a very clear proposal of the experiment I want to do. I send that to a journal, and they send it out for peer review. You do not have to send it to a journal; you could send it to a funder.

Before I start, a collection of experts looks at my experiment and they give me a critique saying, “You need to provide this control and do this additional experiment.” I take that on board and redesign my



HOUSE OF COMMONS

experiment, and they say, "Okay, we think your research question is good enough. Your plan is excellent science. It is interesting. We want to know the answer. You go ahead with it. We guarantee that no matter what results you find we will put it into the public record and we will publish it." It removes my bias—I am not fishing around looking for splashy results—and it removes their bias. The magazines and journals will publish it even if I discover that this drug does not help reduce the size of brain cancer.

That is a big problem we have right now because the academic output is dependent on these companies that run magazines. They want splashy results. They want to put out an article that says, "Chocolate makes you live longer," and all the research that says, "No, it doesn't," doesn't get published as it should. We call that the file drawer problem. All these good experiments go into the file drawer and you never see them, even though we would like them to go out. So, yes, I am a big fan of peer review very early in the process.

Q75 Chair: Thank you very much indeed. You speak with admirable clarity on these things.

Let me turn to Dr Goldacre. To what extent do you agree with the critique that Dr Butler has offered and the direction of her solution to the problem?

Dr Goldacre: I broadly agree. There are two problems with the academic publication system as it stands. One is the challenge of poor information architecture. As Dr Butler says, you cannot see the negative reviews that a paper had previously. You are also often not likely to see the negative results that people have generated when conducting studies if they decide not to submit them.

There is also an additional problem. These days, if you are doing research with data, you are writing code—you are writing software—and the true and complete story of how you produced your results is the code that you wrote attached to the data that you executed it across.

Unfortunately, our model of how we publish scientific findings was built in the 19th century. You write a five-page essay that gets published on paper in a journal. That is a helpful and informative piece of contextual information about what you did, but it is a very long way from a complete story.

I run the DataLab at the University of Oxford, which is where a mixture of software developers, traditional researchers and clinicians build research outputs and research platforms for NHS data. I am very struck by the fact that, when we go to publish our results, sharing the code is viewed as an optional extra. Journals never ask you to share the code. They very rarely review the code. When I review journal articles as a peer reviewer, I often say, "I don't know what they did. They've told me they used a mixed effects logistic regression. That's not enough information for me to



know exactly how they did their study.” When you ask for the code, journals often say, “We don’t really do that. That’s not part of what we do.”

There are well-rehearsed arguments about the problems with peer-reviewed journals in general, but, separately, in addition to that, there is a question of whether peer-reviewed journals have an ambition even to publish the full complete picture of all the work that is done. There are some journals that are now experimenting with new models and requiring people to share code and data, but I think it should be the norm.

Separately, there are things that funders and universities should do to improve the quality of code and code sharing by universities. If journals themselves required it as part of the review process but also as part of their transparency commitment, that would be really powerful.

Q76 Chair: Talking about the publication in journals, is it capable of remedy? As you say, Dr Goldacre, in some ways it is an anachronism that an essay is the only reflection of what is a multilevel study with embedded data and all sorts of other things to produce a final output, which is the one that gets plaudits and is important for careers. Does it require a remodelling and a rethinking of what the output of an academic research project is?

Dr Goldacre: I think it does. I have some views about what a better model would look like. You would have traditional academic journals running in parallel. The ideal system would probably be that all publicly resourced research, or all research done in British universities, would have a project idea attached to it; all of the relevant digital objects such as links to the dataset or the dataset itself; links to the code repository; links to the protocol; links to the final published paper; and links to any additional appendices, time-stamped, attached to data about the funding and so on.

You can picture the perfect information architecture for British academia. The right way to get there would be to resource some groups to start iteratively building prototypes of what that might look like. There are lots of good ingredients dispersed around the system.

A lot of stuff that has been done in the last 20 years has been driven by meeting funders’ and journals’ administrative requirements. We now have ORCID, which is a unique identifier for researchers. There are various projects to try to produce unique IDs for funders. All of this stuff rapidly becomes very geeky—with apologies—because if you want to track activity you have to think about what the unique ID is, the canonical source of those unique identifiers, who manages that, and how they interact with each other. There is a better way, and the way to get there would be to have someone like UKRI resource creative and innovative work by mixed teams of software developers, researchers and policy folk to think through what the optimal model would be and then start road-testing aspects of it.



Related to that, NIHR has done something really quite interesting with its own in-house journals where they do not tend to go as far as good practice around code sharing or data sharing, in part because health data is often confidential, so there are more challenges around sharing the data, but they have their own journal where you have to publish your results within 12 months for certain classes of NIHR-funded research. People publish there in parallel to publication in a traditional academic journal, often with hundreds of pages of detail rather than five or seven, and that is a very successful model. The appendices in there are often very informative and helpful.

Q77 Chair: You say publish in parallel. Can you imagine a world in which reputable and highly respected organisations publish their own work, with confidence reposing in their standards so that people went less to referee journals, on the basis that they can get a more full-spectrum product from the in-house publication?

Dr Goldacre: You want the two in parallel. One of the most useful things that the peer review journal system does, albeit imperfectly, is provide newspaper-like editorial control over not just quality but also general interest. That is a useful role for them to play. Pure self-publication by universities, for example, would create its own different problems, so you want the two in parallel.

The NIHR HTA journal itself has peer review. Some of the things that you want to be shared are digital objects associated with the research project, where peer review is not really relevant at the stage of publishing the results. You want to see the protocol. There is no sense in peer reviewing the protocol that is written at the beginning of the study describing what you are going to do. You do not peer review that at the end; it is just there, it just exists, and you want it shared. It is the same with the code. You might peer review it and say, "I would have done it differently. Do it better. Go away and have another go," but if it is at the point of publication, you just need to see the code.

Chair: Thank you very much indeed.

Q78 Aaron Bell: Thank you both for coming here today. Both of you have already spoken about this to some extent, but I want to talk a little bit more about the incentives that the system creates for you. Could you describe from your own perspective how your scholarly merit has been assessed in your career to date—for example, when you were applying for funding or for a new role? You alluded to this already, Dr Butler. How can the evaluation of researchers better reflect the wider contributions they make to academia? I will start with you, Dr Butler.

Dr Butler: Thank you for the question. I think this is the key question. I am not stating it too plainly and I am not exaggerating when I say that researchers at UK universities are assessed on two metrics: how much grant income we bring to the university and how many papers we publish in an academic journal, preferably in a certain subset—a fancy academic



HOUSE OF COMMONS

journal. That is it. You might have to do a little teaching and you might have to do a little administration, but you are entirely judged by the number of publications in big journals and the amount of money.

This might seem callous at first read, but it makes perfect sense if you are running a university. That is how universities bring in their own income to do research; there is the income from grants, which is fine, and the Research Excellence Framework that the Government use to distribute funds to do research. The amount of funds distributed is entirely mapped to the number of publications in these certain types of journal. There is huge income pressure for us to shape our research to please a certain subset of journals, and to do that quickly.

If you were running a biotech firm and you really wanted to develop a new cancer drug that worked, this is not the way you would plan to assess the success or not of those experiments. It is an understandable method and perhaps a little bit of an archaic method, but it is not the way that you would want to reward research so that you have the most true, the most accurate and the most actionable research.

The other thing is that we are assessed as individuals. Let us say I am a data analyst. I am going to collaborate with a molecular biologist, and we are going to look at protein folding to find a good drug that works on cancer. He and I work in the same department. We find this great drug and we publish the paper. For our promotion documents for the Research Excellence Framework, only one of us gets credit for that paper. That is just the way the bookkeeping works. They are not trying to make us compete, but they are.

If you are in an operating theatre, you do not want your junior and senior surgeon competing for time in surgery; you want the best possible outcome there. In an ideal world, you do not try to evaluate individuals—“Jess got a paper in this fancy journal.” You want to try to evaluate teams.

Ben said how much we need software developers. We work in code now. Universities are not set up to support people like software developers or statisticians who have no ambition—they do not care at all about whether they have a paper in a fancy journal; they want accurate statistics and they want to write code that works. Right now, those people not only do not succeed but they fail. They are pushed out of the system. They could easily work for a company, get paid six times as much and get promoted.

Q79 **Aaron Bell:** We need to find a way to recognise their contribution.

Dr Butler: Yes.

Q80 **Aaron Bell:** From an academic perspective, from your perspective, have you felt that you have been pushed to do things that you would rather not have done and where you would rather have concentrated on something else, because of the incentive structures?



HOUSE OF COMMONS

Dr Butler: All day long. I hate to say it. I probably should not say it on the record, but yes.

Aaron Bell: You have now.

Dr Butler: I will give you a perfect example. I work at a university. I do research. I use health data, but I did not usually work within the NHS. During the start of the pandemic, the NHS needed data analysts right away. They had to build the list of patients they should shield because they were extremely clinically vulnerable to bad outcomes because of pre-existing conditions. They grabbed a bunch of us from the uni and said, "Please get in here and work," and we did. They needed us. I made a little piece of software that mapped the networks of contacts for Test and Trace, which was not complicated. I was able to do it quickly, give it to them, and they still run it every day. It was delightful. It was the most fun I have had at work in ages. I get no academic credit for this because it is not an academic journal publication.

The software that gets used by Public Health Scotland works well. Everyone is happy with it. Teamwork got it out there. I won't get any credit for it until I sit down and try to make it somehow into a magazine article. I just think the incentive structure is too old-fashioned as it stands.

Q81 **Aaron Bell:** Thank you. Dr Goldacre, have you felt those pressures and misaligned incentives in your own career to date? What could we do to change the evaluation of researchers and people working in software development and all the rest of it to better reflect wider contributions?

Dr Goldacre: Yes, I have felt it, and we have learned to hack it in my group. Our interest, as I said, is as a mixed team of software developers, researchers and clinicians. We take large NHS datasets and turn those into traditional academic papers, but we also build live, interactive, data-driven tools. We build software to help other people do their research—research platforms, little nuggets of code that help people do good-quality research. For almost the entirety of my career doing that, I have only ever been able to subsidise that work by getting money from funders that has the object of creating a single research paper. Then I basically steal that money—through fair and legitimate means—and spend it, as far as possible, on doing the work where our heart is, which is building open code, open tools, and open resources to help drive open, reproducible science.

There is almost no open competitive funding available anywhere for anybody who wants to build those kinds of intermediate tools that are intellectually and academically creative and incredibly powerful and high impact in terms of the good that they do. There is almost no open competitive funding to do that kind of work, so that kind of work is not done. It is done by people who are, to a greater or lesser extent perhaps, bloody-minded. It is done almost as a hobbyist activity, as an exhaust



HOUSE OF COMMONS

product of research, but it should be an independent, high-status and legitimately funded activity.

As Dr Butler says, research software engineers are at the heart of this new way of working. When you work with electronic health records data and when you work with any data at all to produce research outputs, you are writing code and software. You need people who are really good at writing software—not people who are researchers who have dabbled a little bit here and there. You need people who are professional software engineers on your team.

The Society of Research Software Engineering and the research software engineering movement got going in the UK about 10 years ago. It has been quite penetrant in some quite small areas. You will see one group in one university who do a lot of stuff on digital archaeology maybe, but it has not become mainstream, in large part because of two things.

One is that funders do not really recognise it, understand it and fund it. They are trapped in a model of, “Everybody who wants to work in academia is desperate to be an academic.” I employ software developers whose outside option is earning very good, low six-figure salaries. They are willing to halve their salary to work with me if it is an interesting project. They are not willing to decimate it to £30,000 or £40,000. Universities in turn do not really understand that a full-stack commercial-grade software developer is a different set of skills to the person who fixes your printer when it is broken. They think that the pay grades should be defined by, “You’re an IT person. We have a salary grade for that. How many people do you manage?”—“None. I would never waste her on managing other people. That would be a disaster.” And they go, “Well, in that case, 42 grand a year.” That just does not work.

Also, the funders do not understand that we are trying to seduce people in to do this work. When I speak with funders, they will say, “Perhaps we could have a research software engineering fellowship scheme for the lucky people who desperately scrabble to get in through the door.” They will have to write out a 40-page form with lots of information about how desperate they are and their previous track record in academia, when they will have none, but they will have built the entire retail back end of Marks & Spencer or something like that. We need funders and universities to engage not in ideological or narrative terms with how to get great software researchers working in parallel with researchers, but in concrete terms.

The best way to do it would be through some pilot schemes where people say we are going to have a big research software engineering squad in this electronic health records research group, this NHS data analysis group, this digital archaeology group, or whatever it is, and have them as pioneers showing the value of modern, open, collaborative approaches to computational data science, teaching people how to code and share their code, and how to make efficient code.



Where that happens at the moment, you should never brush over it. There are things like the Software Sustainability Institute, the NHS-R Community and the Software Carpentry movement, but they are all hobbyist activities. It is mostly people's part-time hobby and it should be mainstream, foreground activity funded by UKRI and NIHR.

Q82 Aaron Bell: Thank you, Dr Goldacre. I was a software developer before I came here, so I recognise some of what you have just been saying.

If I could turn to the reproducibility element of our inquiry and the incentives there, could you describe the impact that the work conditions, such as the short-term contracts and the pressure to publish, have on the research output for early-career researchers in particular? Are they incentivised to produce open and reproducible research, or are they incentivised to produce something that is interesting rather than necessarily reproducible?

Dr Goldacre: A lot of that is irreducible in some senses, and we have to be a bit realistic about a competitive environment and low-resource environment. People are incentivised to do, to an extent, not necessarily p-hacking but to select for interesting things and find interesting things to publish, which may not give a full reflection of all the work they have done.

The bigger problem is that the single-minded focus on getting any paper, anywhere, as good as possible, and it being an ongoing race, means that all the other stuff around documenting your code, making beautiful code, and turning your one-off script into a reusable function that other people could use in their subsequent research is so massively pushed down the priority list that it does not happen. I think that is the real challenge.

Q83 Aaron Bell: Thank you. Dr Butler, what are your thoughts about the incentives for early-career researchers in particular? Are they incentivised to produce top-quality, open, reproducible research by the system we currently have?

Dr Butler: No, they are actively disincentivised from producing top-quality, reproducible research. It is true and it is worth stating plainly. It is worth knowing that the vast majority of research done by the UK is done by people who are hired on temporary contracts to universities for about a year. Let us say I am a professor and I receive funding to do some cancer research—it is a two-year grant, which is entirely typical—to figure out if this drug shrinks brain tumours in mice. I have to get permission to hire a researcher on that grant. I have to teach. I have to do other things. I put out a job ad saying, "Would you please come to Aberdeen for 11 months or 16 months to do this research for £30,000 a year? You need to have a PhD."

They come in. Let us say in the best-case scenario I could get this researcher trained up to understand all the maths stuff, all the chemistry, where everything is and how the university works in three months. Almost as soon as they get started on this research, not only does their



contract end but they are automatically made redundant. There is no controversy there. They lose their visas. When they start this project, they are looking for their next job. They are looking for that next grant. They are looking to go out to industry. Even if they are not looking, and they decide they would like to work purely for one year, the only thing they get judged on at the next step is how fancy the magazine is in which they published the summary of their research.

No peer reviewer—no one—will look at the code they wrote. They will not look at the data. In all my years of research, no peer reviewer has ever asked to look at the data behind my experiments and the code. It is unjust to hold these people on these precarious temporary contracts to this really high-quality standard when you are judging them by something else entirely and then you are cutting off their position. That is not an exaggeration; it is just how it works every day.

Q84 Aaron Bell: How should it work? How should leaders in research institutions like universities ensure that academics at all stages have the time and the incentive structure to conduct that top-quality research we all want to see?

Dr Butler: We need to recognise that not everyone is a principal investigator and a professor at Oxford, and that research should be done by teams. You need a big figurehead to lead the research and have the vision. You also need workers. You need someone who is a meticulous chemist. You need a careful statistician. You need someone who is paid and judged by how good their code is and how well commented it is. Then you need to judge the entire output of that team by whether or not the experiment worked. Give it to another team. I say I shrunk the mouse tumours by 85%. Did I? Go look. Reward me if I do and reward the team that is testing that. It is a paradigm shift. We will have to slow down. We cannot write three papers a year. We will have to recognise that there are full-time positions in quality assurance and in software engineering at the UK research institutes that happen to be universities. Slowing down and recognising that we care about quality is what we have to do.

Aaron Bell: Thank you very much. Thank you both.

Q85 Chair: Thank you, Aaron. Before turning to Rebecca Long Bailey, I want to follow up Aaron's line of questioning to Dr Goldacre. Both of you have described very clearly a problem that funders are focused on certain outputs such as articles in refereed journals of prestige rather than the greater spread of the products. When we think about funders, there are medical research charities and institutions like the Wellcome foundation, but much of it is publicly funded and is through UKRI and, in particular, the research councils. The research councils consist of academics in universities. It is self-run and self-regulated. It is the nearest we have, in many respects, in this country to a co-operative enterprise.

The failures that you are describing are the failures of your colleagues,



HOUSE OF COMMONS

are they not? This ultimately derives from the Haldane principle that says that when public money is provided it should be down to researchers to decide where it goes. Does this say that that system does not work and there should be a more directive approach taken by Government and public bodies to the conditions in which grants are allocated, because academics are not doing a good enough job?

Dr Goldacre: I do not think it is one or the other. Often, it is necessary, if you want to drive culture shift, to take a slightly muscular hand to shifting the direction of how things work. When it comes to reproducible research, the future is already here; it is just very badly and imperfectly distributed. There are some groups in some areas and some specialties where you get really good collaborative work between software developers and researchers, where there is much more of a norm around people sharing code and data. Structural biology, structural genomics and a lot of physics is very much built around team science, sharing code and sharing data.

Work with NHS data is particularly far behind, and that is a particular structural and strategic failing. Funders could do a lot by setting norms in very practical ways. They could say, "We require that you share your code and results." They could enforce that by checking that you have done it, before making the final payment. They could check that you have done it in the past by looking at previous outputs or making the applicant sign a thing saying, "I have always shared my code and my results," before they make new awards. That would make a big difference.

It is also important to recognise that this problem goes beyond pure academia. We published papers in *Nature* and in *The Lancet* on the relationship between ethnicity and death from Covid-19. We did that in our research platform, OpenSAFELY, which we built with the NHS, which imposes open working methods by default. You can only run code across 58 million patients' GP records in OpenSAFELY by first putting it on GitHub, which is a place where people share their code and software. Everything that runs on that is done in a fully open and reproducible way, and it imposes open working on the whole community.

In the same time period, there were reports published by ONS and PHE looking at the relationship between ethnicity and death from Covid-19. Each of those three studies—ours, ONS and PHE—gave slightly different results. I do not think that any of them were dramatically wrong, but when you are trying to angle in on the truth, you are trying to triangulate the truth and you are trying to get the most accurate answer, you want to know exactly what was done. You want to understand why ONS, PHE and OpenSAFELY got slightly different results for the relationship between being black or Asian and dying of Covid. It could be a difference in the source data. It could be a difference in how the source data was prepared into an analysis-ready dataset. It could be a difference in the statistical model. To understand how that was done, you need the code for all of the different analyses. For ONS and PHE, there isn't code in the open.



HOUSE OF COMMONS

That is not because it was done by bad people; it is because that is not a norm. It is not just a failure to have that norm in academia; it is across the board.

In addition to that, ONS, in particular, and PHE, in other parts of their activity, have led the way in sharing code. Public Health England have a fantastic Covid dashboard where they share all the code on GitHub. ONS have developed something called the reproducible analytical pipelines project, which is a set of working practices for open working and technical documentation when producing Government statistics. They have a training programme associated with it, and a set of norms and house styles. It is an incredibly powerful and incredibly valuable project that covers a fairly small but growing percentage of all Government statistical work.

This problem goes beyond academia, but it is also being fixed in bits of academia, in bits of Government stats and in bits of the health service. We need to make sure that we spread those shoots across the whole system.

Chair: Thank you.

Q86 **Rebecca Long Bailey:** Thank you both for speaking to us today. Very generally, to what extent can individual researchers produce work that is completely open and reproducible if they wish? I will start with Dr Butler.

Dr Butler: It is easy. The infrastructure exists. I can write my research protocol before I start collecting data. I can pop it on to Open Science, and you can look at it one minute later. If legally possible, I can make all of my data open. The digital infrastructure exists for me to work this way. I can request that peer reviews be made publicly available by choosing a certain type of publishing platform that works that way.

It is entirely possible to work completely openly from start to finish. I can publish my output on the internet without going through a journal. I could allow you to alter some of the assumptions I make in the work I do to see if that changes my result. I could request colleagues to try to reproduce my result—let us say they could say yes, given their workloads. The infrastructure exists. I would not recommend it to anyone trying to get ahead and get rapidly promoted, but it is entirely possible and I would even say not that difficult.

Q87 **Rebecca Long Bailey:** Thank you. Dr Goldacre.

Dr Goldacre: You certainly could, in the sense that the platforms are there, but you may not have the skills or the time. My group published an editorial in the *BMJ* a year or two ago titled "Why epidemiologists should share their analytic code". This was after a very prominent randomised control trial had been retracted because a coding error had led to them getting precisely the opposite result to the true one. They flipped the variable. You could have seen it if you had been able to look at the code.



We discussed it with the editors at the *BMJ* and lots of researchers, and the most common response we got from epidemiology researchers was, "What are you talking about? We don't write code. We just write statistical analysis scripts in R and Stata." They do write code; it is just that they write code without the framework, skills or mindset of a software developer. They think, "I'm just writing a script." They do not even have the words for what they are doing.

In the absence of good software carpentry skills as well as the absence of incentives, it is very hard for people to do the right thing. We need to train them and incentivise them, including with things like the REF, which has not come up yet but is the standard way that all universities are judged. That is basically papers and money for the most part, and then some notional content about things like environment. The environment statement in the REF should include an expectation that you have support on software development rather than just doing it in an ad hoc way.

Lastly, you can submit a GitHub repository full of code to the REF if you want to, but almost nobody ever does because the reviewers on the panels would not necessarily know how to evaluate it and it would not be regarded as a high-status research output. The REF could fix that by saying, "You do one stage of research by writing code. You have submitted lots of quantitative research papers where you must have written code. Therefore, in the REF, for every 10 research papers that are submitted doing quantitative research, we expect to see a minimum of two GitHub repositories with code attached to those, otherwise they are not admissible."

Q88 **Rebecca Long Bailey:** Thanks, that is really helpful. Is there anything else that you think research institutions can do to incentivise the open and reproducible methodologies that a researcher could use, Dr Goldacre? You have already touched on a few ideas already.

Dr Goldacre: It all comes down to training in research software engineering and employing research software engineers to work alongside those research groups. You need to have the incentives, but you also need to have the skills.

Q89 **Rebecca Long Bailey:** Thank you. Dr Butler.

Dr Butler: I think universities could. You could request that promotion criteria include evidence. There is a great paper by Danielle Rice that came out last year in the *British Medical Journal*. Her team looked at the promotion criteria of 170 medical schools from top universities. All of the schools require a list of publications and what journals they are in, and how much grant you have been funded. None requests any evidence of open data or code, studies that have been reproduced or application. This is a universal problem.

My request for the Committee is a recommendation further up. As soon as UKRI said all publicly funded research has to be made available open



HOUSE OF COMMONS

access at the end, we did it. It was not easy, but we did it, and we did it quickly. It is morally right. But I feel that that simple pressure from above will help.

The best score you can get on the Research Excellence Framework that brings your university the most money is summarised by the phrase, "This paper was world leading in its originality, significance and rigour." Is that what we want? Is that the best possible research that comes out of the UK? That phrase could be tweaked. Everything has to be original and significant all the time. Do we want to be a little more specific about how we measure rigour, or do you just want to say UKRI-funded research has to have publicly available methods and data openly available as far as possible? These tiny requests could have huge impacts.

Q90 **Rebecca Long Bailey:** You have both mentioned the restrictions you face in terms of reproducibility but also the restrictions that you both face in your research overall due to the way that research is funded and because you have to prioritise certain types of research to get into the big, glossy magazines. Do you think there is a grassroots movement building now within research and academia to push for change? How can research institutions themselves support that supposed grassroots movement? Dr Goldacre first.

Dr Goldacre: It all comes down to practicalities. There is a grassroots movement, but, like I said, it is a hobbyist activity. It is not a high-status, independently funded activity. The last thing that is really encouraging to see, in part because of work from this Committee on trial reporting, is that there is a bit of a shift. To talk about reproducibility in research and clinical trial results not being published, for example, was regarded 10 years ago as being quite transgressive or radical, whereas now it is increasingly more mainstream and acceptable. That is a consequence of institutions being built around it. I do not think culture change comes from words. Culture change comes from where the money flows. If you resource research software engineers and make funding contingent on modern, open, collaborative approaches to computational data science and code sharing, the behaviours will follow. It is as simple as that.

Q91 **Rebecca Long Bailey:** Thank you. Dr Butler.

Dr Butler: There is a very active grassroots movement. If I define that as people who are more junior in the system, I agree. There is an army of people who would prefer to work in that more rigorous way to produce better science, to have the science applied broadly and to care less about splashy headlines. There is an absolute army of people who would like to work this way.

They are primarily young and often on temporary contracts. They are at one university for a year and another university for two years. Then they get a better offer from industry and they leave. The Committee could support them by offering something like accolades. Endow a position for



HOUSE OF COMMONS

this type of work. Give it to young people or just award it something that shows to those higher up that what they call metascience—how we do better science—is worth while and needs to be rewarded.

Rebecca Long Bailey: Thank you. That is really helpful.

Chair: Thank you very much indeed.

Q92 **Graham Stringer:** This question is to Dr Butler. If the codes and data are not examined and available, that simply means that the research is not reproducible, does it not?

Dr Butler: It could be reproducible. Maybe the data is perfectly tidy and the code is perfectly written, but it is not being reproduced.

Q93 **Graham Stringer:** Let me ask a related question. When the predecessor Committee of this Committee looked at climate research at the University of East Anglia, we were surprised to hear, as you have described, that the codes and data were not looked at by peer review, but also, because some of the data came from meteorological stations all over the world, the University of East Anglia did not think that they had the legal authority to publish it. Is the legal ownership of data in these datasets a problem?

Dr Butler: Sure. People are often frightened, whether it is justified or not, that they have the right to let go of this data. Reassurance that this should be the norm would go quite far. I work with patient data. It legally cannot be made open. That is fine. There are high-security servers we could put it on and only allow access to peer reviewers. This is a surmountable problem. Making it the norm to make data open is important.

My biggest concern coming in front of this Committee and being very frank about the problems I think we have with research rigour around the world is that it would be used against us. My biggest worry is that me saying we could work to a higher standard will be interpreted by some as, "Coronavirus does not exist," or, "The climate is not changing." It is terrifying to me, frankly, but I think we have to be frank about it. Being as transparent as possible is the best defence we have against people who say this research is not correct. Normalising this openness is incredibly important.

Q94 **Graham Stringer:** We have heard in the oral evidence during this inquiry that, when eminent scientists are reviewing grants, Nobel prize winners often say, "If I had had to have this short-term look at things, I could never have done the research that led to my breakthrough in" whatever scientific fields they were in. If one looks at the major scientific breakthroughs of the 20th century like DNA, to use an obvious example, Crick did not do anything for 15 years, and then with Watson he made the great breakthrough on the structure of DNA. Do you think that those major breakthroughs are more difficult given the funding regime and the incentives within the research community?



Dr Butler: I do. Darwin was a gentleman intellect who was able to take 20 years to write one book. We are not able to sit and think deeply for several years. We would lose our jobs, frankly. I recognise the Government cannot support ivory towers full of people drinking coffee while looking out their windows and thinking deeply, but I think we need a balance. We need less research and better research. We need time for criticism and rigour. So, yes, I highly recommend slowing the output system down.

Graham Stringer: Thank you.

Q95 **Chair:** Dr Butler, we are very grateful for your candour, particularly what you said about it sometimes having consequences. I hope those consequences will be reputationally positive and it is admirable that you are willing to appear and be so candid.

Right at the beginning of your evidence to us you mentioned registered reports, which is to say that at the beginning of a piece of research work you lodge your hypotheses, your intended methodologies and the rest of it. One can see the argument for that. Is this suitable for an age that is closing in terms of how we might do research? We now have access to very large datasets and, through machine learning, have the ability to spot connections between variables that perhaps a human has not thought, in prospect, have a connection. If we limit ourselves to following what a human has designed as a research project, are we not cutting ourselves off from a whole world of quite impactful discovery, particularly in medical sciences, as a result of this kind of approach to rigour and regulation?

Dr Butler: Not at all. The good news is that you could easily design an experiment that uses machine learning on massively big data and do it as a registered report. All it needs is for you to sit down with your peers—the ones you are working with and, hopefully, peer reviewers who are critical—and say, “I would like to do machine learning on all of the GP records in England to determine the connection between weight and death.” Make it up; it doesn’t matter. You say, “Here is how I will determine if the data is of good quality.” You cannot have a good machine-learning analysis on data where almost all of the records for weight are empty. You have to go through and say, “Are there biases in the data of the set I am using? Perhaps nobody over the age of 60 made it into this dataset.” You need to think through that quality control process before you go fishing around for a result.

You can easily pre-register decisions. If it looks like this, we will come to this conclusion. If it looks like this, we will come to that. Maybe it locks you in a little bit. You cannot fish around for data, but you could easily pre-register a machine-learning study and get those criticisms of the study of the dataset early on, which are incredibly important for that type of work, and then you get the guarantee that your result gets out to the public record. Maybe you do this excellent analysis, and you are looking



hard for any type of connection, but you do not find anything. That can still go into the public record if you publish by this format.

Q96 **Chair:** If you are looking, for example, at a connection between weight and the incidence of Covid, and you do not discover anything particularly interesting from that, but you discover something, to take your example, between weights and deaths more generally in certain parts of the country or certain minorities, which may not have been at all what you were looking for—you were looking specifically at Covid, but this is thrown up—does that research design and the registration of your intended approach not constrain your ability to make use of that finding?

Dr Butler: No, not at all. In a registered report I did, we found some incidental findings that were interesting. All that is important is that you are incredibly clear in the record that it was an incidental finding.

Chair: I see.

Dr Butler: You say, “This is what we proposed to do. These were the outcomes. While we were doing it, we found this incredibly useful and interesting connection that we are reporting here to you. Go off and follow up on it, and see if it holds another dataset.” All we are asking for is transparency so that when you are reading my report of it you know how I approached the problem that I found accidentally, and you keep that in mind when you design the experiment where you go off and test it more rigorously.

Q97 **Chair:** Thank you, that is very clear. Dr Goldacre, do you have anything to add to what Dr Butler said?

Dr Goldacre: No, I fully agree. There is hypothesis-driven research and there is exploratory research, and you need to clearly state which it is that you are doing and where your finding came from.

Chair: Excellent. Thank you very much indeed again for appearing in front of us. You have been extremely informative. We have slightly overrun our time, which is an indication of that. I will say goodbye to our first two witnesses, Dr Goldacre and Dr Butler.

Examination of witnesses

Witnesses: Dr Dhand and Dr Moylan.

Q98 **Chair:** I welcome our next set of witnesses, who come from the academic research publishing sector. I am very pleased to welcome Dr Ritu Dhand, the chief scientific officer of Springer Nature, and Dr Elizabeth Moylan, a publisher in research integrity and publishing ethics at the publisher Wiley. Thank you very much, both of you, for appearing before us today.

Perhaps I could start with a question to Dr Dhand. I think you caught some of the earlier session looking at the role that universities and research institutions play in upholding research integrity. Perhaps I can



HOUSE OF COMMONS

ask the obvious question: what role do academic journals and publishers play in upholding research integrity?

Dr Dhand: Thank you for the invitation. Yes, I listened, and I sympathised with and recognised a lot of what was being said. Publishers are doing quite a lot to look at and uphold reproducibility and support open science, which is a big part of reproducibility. We have checklists to check on statistical reporting. We have checklists to look at where data exists. Is it in an open repository? We encourage and, in some cases, mandate that data is put into repositories. We have an open repository—a deposition of protocols. We encourage, and in some cases mandate, that these very detailed protocols are reported.

We support open science and therefore we absolutely support pre-prints, open data deposition and open protocols, but we do not always mandate these things. That is the difference.

We support community norms. Dr Butler and Dr Goldacre mentioned communities like genomicists and structural biologists where the norm is to deposit into open datasets. We police those norms. We do not encourage and we do not just mandate: we actually police that norm.

Q99 **Chair:** Dr Dhand, what would you say to the suggestion that was made in the previous panel that—not through any fault of your own—structurally, the referee process for a journal is at the end of the research journey, by which time it is too late to have any influence on the design, other than a binary decision whether to publish or not, and therefore the whole thing needs to be front-end-loaded and rest more with universities?

Dr Dhand: I completely agree. We are the end of the process. We are catching research after it has been completed. What you really want is researchers to do research with these design points and these factors in consideration.

To give you one example, we have a reporting summary that asks for a whole bunch of checks on things that should be reported but not all researchers do report, especially in the biomedical sciences—things like whether your study was randomised and blinded. If you were doing studies with human patients, that community works to a higher standard of transparent reporting about their statistical analyses, and you would get those things included. If you are working with mice, which is what most biomedical communities work with as a model organism, or worms or flies, the standards of reporting are not as robust. Therefore, it is not the norm to report whether a study was blinded or randomised.

We think those kinds of things should be reported, so we ask researchers to fill in a checklist and make known whether it was or was not. We are asking for transparency. We go through a whole bunch of questions like that about the experimental design that not everyone will report as the norm.

Chair: I see.



Dr Dhand: Researchers have come back to us and said, “I think it was a pain, but I’m really glad I’ve done it. Now, I am asking my postdocs and PhDs to think about all of those things when they are designing the experiments, as that is when it should be happening.”

Q100 **Chair:** Thank you. Perhaps I can turn to Dr Moylan. On the question about involvement earlier than at the final point of publication, one of the suggestions that was made is that journals, for reasons that are understandable, are more likely to publish results that are interesting, which means that results that have been very faithfully and rigorously researched and that confirm an existing hypothesis, or fail to support a new one, simply do not see the light of day, and that gives a bias in what is published. Is that an inevitable feature of a journal? I realise there is a range of popularity and sales associated with different journals, but do they share a common denominator in that their readers want to see what is interesting and new rather than what may turn out not to be?

Dr Moylan: I do not think that is a problem of the journal; that is a problem of how the researchers perceive things. Often, they will want to show the positive stories and not necessarily report on something that was a bit confirmatory. Journals would welcome that.

Q101 **Chair:** You are as likely to report a paper that is confirmatory as one that is new.

Dr Moylan: Yes, because it validates and reproduces that previous study. There is a variation across journals. There are journals that are absolutely geared up to publish sound science whatever the outcome. We would welcome that and we would share that. There is a problem of researchers getting round to sharing that, which is the file drawer problem that has been referred to, but we would welcome publication of those results.

Q102 **Chair:** You have a portfolio that deals with research integrity and publishing ethics. Do you have any statistics to bear out the interesting observation that you made that you are as likely to publish something that is confirmatory as something that is new and surprising?

Dr Moylan: We would have to check what each journal was publishing. We publish over 1,600 journals. It would take a lot of effort to show that in a particular journal—

Q103 **Chair:** We are talking about research. When one has a large number of observations, one has methods for doing that. That is not a trivial question.

Dr Moylan: I think what you are asking is within the journals that we publish how many of those articles are publishing confirmatory sound science results as opposed to exciting results. Is that the question you are asking?

Q104 **Chair:** You made a very interesting observation. What we heard from the



previous panel was that journals were much more interested in publishing research results that were bold in what they found and surprising to some extent. You told the Committee that that is not the case, and that you are as likely to publish confirmatory papers as non-confirmatory ones, so I am asking for the evidence behind your very interesting assertion.

Dr Moylan: It is to do with the journals and their scope. If they are willing to publish sound science articles in that area, they will.

Q105 **Chair:** I am not asking you about their intention; I am asking whether you have evidence to back up the assertion that you have made.

Dr Moylan: Yes, I believe we will have that, and we can share that with the Committee. I will have to talk to my colleagues to show that we publish sound science results, but it has to come from the researchers. It is not the publishers saying that we will not publish that.

Chair: Indeed. I will turn to my colleagues.

Q106 **Graham Stringer:** How easy is it for the publishers to put all the data that underpins the research papers that are published online?

Dr Moylan: If a researcher is operating with open research practices and they are sharing their data and depositing it on an open repository, within a research article the publisher will publish a data availability statement and link to that repository. We do not have to hold all that data. We would link to where that data resides.

Q107 **Graham Stringer:** Is that common practice?

Dr Moylan: Yes. As we have talked about, it is a little dependent on the research communities and what stage they are at in their open research journey. Some journals will encourage data sharing, some will expect to say something about where the data is, and some will mandate the sharing of that data. We have over 800 journals at Wiley where we expect to see a data availability statement saying something about where the data is.

Q108 **Graham Stringer:** Thank you. Going back to a similar point that Greg made earlier, what percentage of publishers offer registered report formats?

Dr Moylan: Registered reports is a growing initiative. It is definitely being offered by many publishers. It is also a little bit subject-specific. Where it works well in fields like psychology, it would be offered in psychology journals across publishers. The Centre for Open Science keeps a list of all the journals that offer registered reports. We can get that information for the Committee.

Q109 **Graham Stringer:** What incentives do publishers have to publish replicable and reproducible research generally? Are there any incentives in the system to do that?



HOUSE OF COMMONS

Dr Moylan: We want to publish research. We are geared up to publishing research and making that available. That is absolutely something we want to do. So I would encourage that, absolutely.

Graham Stringer: Thank you.

Q110 **Chair:** I want to probe that a little. You have a stable of publications. Some are bought largely by libraries and institutions, some with a broader subscriber base. Certainly, for the latter, surely to pique the interest of subscribers, there must be more of an incentive for you to publish more striking findings than ones that confirm but do not add to or advance the stock of human knowledge.

Dr Moylan: I would have to talk to my colleagues who are more into the various partnerships we have with subscribers and the way the publishing landscape is changing now. It is moving from a subscription economy to a more open access-based business model. I have not thought about that aspect. When people subscribe to the journals, they have a subscription to read what is published, so they have access to all the content.

Q111 **Chair:** Dr Dhand, on this point, the Committee has heard representations from the publishing industry as to how important it is that it is maintained and continues to prosper, and it is absolutely foundational to the reputation of UK science that it should operate like that. It is coming under quite some degree of criticism from those who say a more open access model should be practised.

These questions are, it seems to me, of existential importance to the industry. What response do you, as an important player in the industry, have to these questions of whether the incentives on publishers have the effect of biasing or causing a set of incentives to take place among researchers that is not in the collective interest of the country?

Dr Dhand: I think you are asking whether, because of the open access way of publishing, journals would publish science that may not be as robust because each paper has a monetary value. Is that what you are asking?

Q112 **Chair:** No, I am not saying that. I am saying something different. There is no question that you do not publish things that are robust. It would seem to me, and it has been put to this Committee, that as a publisher you are more likely to publish things that have striking and fascinating results rather than the results of studies that confirm what we knew before.

Dr Dhand: First, there are very, very few very, very striking results, or things that change the way the landscape may work in science. Most results build incrementally on something that has been published. That incremental increase is what different journals are interested in. If you are a very specialised journal, the smallest increase of knowledge in any field is of interest. When I was at the bench, if I identified another tyrosine that was phosphorylated, which might sound very boring to most



people, I was very excited. But not everybody will be excited by that because I have not cured cancer; I have just found another tyrosine that is phosphorylated. That means a lot to me because I know what that will mean one day when I build upon this. I will publish that, and there will be a journal that is interested in publishing that, but it is quite a specialised and incremental finding.

Most studies confirm science by having some incremental increase in knowledge. No scientist wants to sit there and just reproduce what someone else has done. Science is intellectual and stimulatory. You want to learn more. You want to learn something different. Most science builds upon what has been published. What journals are then publishing, or looking for, is an incremental increase in knowledge. Different journals will have different interests. Dr Goldacre referred to it in this way. Some journals have a larger audience where, quite frankly, no one will understand the relevance of why another tyrosine phosphorylated on this molecule is interesting. The audience is too broad.

However, a cancer journal looking at my specific type of cancer and my specific type of molecules will be just as excited as me. That is where I can submit the work to the journal. It is how big a leap in knowledge you make and which journal will be interested. Dr Moylan talked about sound science journals. Sound science journals are not interested in any large leap. They will publish the most incremental finding that builds on a field. They confirm the results and they build on results, but the increase is incremental.

Q113 **Chair:** I am grateful for that. Before I turn to Katherine Fletcher, you made the point that there was no interest on the part of a researcher in literally replicating and simply confirming that the findings and approach of a previous study were sound. There is a public interest in that, is there not? It is useful to know, but it may not be career advancing for an individual to devote themselves to confirming that.

Dr Dhand: Would scientists do that? They have to be incentivised to do that. Is it worth funding bodies to do a repeat? Dr Goldacre said at the very beginning, "I just have to write a proposal for a research project that they want to do." If they say they are going to repeat what someone else has done, it is highly unlikely they will get funded, because a funding body also wants to see science advance.

Reproducibility also comes from building on science. You do not have to do exactly the same research to know that previous research was true. If you have an incremental build on it, you still verify existing research. The way a research paper is written is that you note all the incremental parts of the story that have been published everywhere. I have confirmed this because I have got this far; I have confirmed *this* because I have got *this* far. Most papers report 75 to 100 references that usually refer to what is in the literature, and then they say, "And this is my new bit." They do confirm science.



Chair: Let me turn to Katherine Fletcher, who is certainly interested in these detailed advances.

Q114 **Katherine Fletcher:** Quality nerdery on a tyrosine! If you will forgive me, I am going to pick up the tyrosine phosphorylation analogy, but, before, I need to put on the record that the next few sentences I come out with are not true. What we are interested in here is the idea that somebody can, through a set of perverse incentives, or maybe wanting to make sure they continue to be funded, report something as true that is not true, and then the builds on it end up going down a blind alley. For example, your discovery of another phosphorylating tyrosine is brilliant, but you are doing that because there is a paper somewhere that says that tyrosine phosphorylation is important in cancer. If, for example, that was wrong, your incremental approach is based on something that is fundamentally incorrect.

I will start with Dr Dhand and then I will come to you, Dr Moylan. What processes are you putting in place to facilitate constructive criticism of individual papers? In the simplest terms, how do we make sure that we are not building on sand as foundations?

Dr Dhand: That is a very, very good question. Given the way scientific is frameworked, when I find that tyrosine was phosphorylated—I looked at the papers before to get here—and try to get as far as what was already published, I should hit a roadblock if the paper that I need to build on is not true, because I must be able to repeat the research. Given the processes that publishers have in place, that researcher will write to that journal and say, “I tried to do this. I needed to do it because I want to go further, but I have to get past this step and I cannot repeat it.”

Q115 **Katherine Fletcher:** Do you publish those letters?

Dr Dhand: Yes. Journals call them different things. We call them “matters arising in our journals”; other people call them correspondence. I do not know what everyone calls their criticisms of published papers. I also honestly cannot tell you that everyone does this, but we do, and I know many journals and publishers do because it is so important.

When we get criticism, we go through a very a rigorous peer review process. We have to work out whether we are looking for a correction here? Is the person who is complaining at fault? Did they do something wrong? Are we looking at a correction where it is not the case that the whole research is no longer true but a small part of it might need to be corrected? It requires a more rigorous peer review process than an ordinary peer review process where you have a fresh piece of research.

Q116 **Katherine Fletcher:** Is that peer review process published by the end of it, or is it done behind closed doors?

Dr Dhand: That depends on whether the journals have open and transparent peer review—that is, they will publish reports. If they publish reports of peer reviews, then it could be. Most journals at this time do not



have transparent peer review. For example, we have journals that have transparent peer review, and we have journals that are about to roll out transparent peer review, but we also have journals that do not do transparent peer review because not everybody agrees with it—

Q117 **Katherine Fletcher:** Just let me chase to the end and then I want to bring in Dr Moylan—apologies for talking over you. If that peer review says, “This is nonsense. The guy was smoking his socks when he wrote the paper,” how easy is it to retract it so that people are not building on a house of sand in that incremental approach that you have described so brilliantly?

Dr Dhand: If we are absolutely sure that that paper is not solid or sound and no longer conclusive, we retract it.

Q118 **Katherine Fletcher:** That is obvious. Do you think that is general practice in the industry across all the different publishers, or is it down to individual journals?

Dr Dhand: I think all journals are gatekeepers of scientific literature. You cannot keep it a secret. If something is not true and someone has come and told us that, and we ignore them, somebody else will find exactly the same thing. That is how science works. You cannot hide it or pretend it is not true. The scientific community starts talking about it and the problem will get bigger and better known. You cannot walk away from a paper that the community has learned is not true.

Q119 **Katherine Fletcher:** It is just a question of timing. Does it take two minutes, two years or 20 years, but it will be found out in the end?

Dr Dhand: It takes a long time, because you have to be sure. Like a police investigation, what do you need to take that person to court? A little bit of evidence won’t cut it.

Q120 **Katherine Fletcher:** Dr Moylan, what is your view on those questions? I will not repeat them to save time.

Dr Moylan: I entirely agree with Dr Dhand. I would just add that all journals are members of COPE, the Committee on Publication Ethics. There are standards and guidance we adhere to, and when these kinds of issues are raised we as publishers have a duty of care to investigate and follow up. It may not necessarily be a bad thing. It might be an honest error or mistake, made for whatever reason, that did not work. That plays out either in letters to the editor or in a correction to the article, but if it is fundamentally flawed it might play out as a retraction.

Q121 **Katherine Fletcher:** The thrust of both of your learned attentions is that there will not be malign action in terms of mistakes made. One of the things we are considering is the possibility of deliberate tweaking or worse going on. Is there a system to check and enforce for that? Do you think there is any appetite within publishers, along with Government, to create a framework to try to look for those very rare—I grant you—but



HOUSE OF COMMONS

potentially disastrous cases of research misconduct? Dr Moylan, do you think the publishers will be interested in that?

Dr Moylan: Because we have a duty of care towards the integrity of the literature, absolutely. We would follow up and investigate. Those sorts of questionable research practices are often quite tricky to spot and it might take a few years for the actual issue or what happened to come out, but we will follow up.

Dr Dhand: I absolutely agree with what Dr Moylan has said. We do do things now. We have plagiarism checks on all our papers to help us spot these things. We do image integrity checks, but not on all our papers because we do not have it rolled out for all of them. But all of these things are harder to spot at the end of the process. The kinds of things you can spot in a published paper are not as informed or as great as you would spot much earlier in the process.

It was either Dr Goldacre or Dr Butler who said, "Who looks at the raw data when research is being done?" I would say it should be the PI, the group leader, looking at the image that will be published in the paper. Somebody should be assessing the raw data that will become that published image or process paper, but not all PIs have time and that is a huge problem. The more successful you become, the less time you have to do those things in the lab. A new PI will be looking at raw data. They have more time, although they are also writing a lot of grants. They spend a lot of time writing grants. Successful PIs are travelling; they are sitting on committees; they are peer reviewing. They are doing everything else but spending a lot of time with their postdocs actually looking at the data that is produced.

Another problem with the way research is done is that we have an academic framework where we think it will all be okay. If you are in industry and you are a person who will be travelling a lot and will not be managing your group, you will put a framework in place with a deputy who will then take on the responsibility of whatever the PI did. It is expected to happen organically. There will be a postdoc who steps up. There is no position or recognition of that; in fact, you are seen as a failed scientist as a postdoc who stays in a lab to support their PI and do the things they cannot do because they have to do a lot of other things.

Q122 **Katherine Fletcher:** Forgive me, it is such a long time since I was in a lab. I cannot remember, but does a PI always get named on the paper publication?

Dr Dhand: They do if the research was done in their lab, but not the postdoc who is just checking and giving advice.

Q123 **Katherine Fletcher:** Essentially, there is a key role that is not necessarily transparent to those reading it or that, to your point, is enforced. Thank you; I think that is really helpful. In a way, that moves me towards where I want to go.



What should the industry be doing to try to support attempts to make research openly accessible, reproducible and replicable, and for fraud prevention? It sounds as if there is something in there about PI supervision and a more formalised structure of raw data checking. You are quite right that, once the graph has been produced, if it has not been monkeyed with, it is only as good as the numbers in the Excel spreadsheet that produced it, or whatever it is. Dr Moylan, what action do you think the industry should be taking to make sure we can have confidence in published research and that it is reproducible?

Dr Moylan: I think all stakeholders have a role to play, and open research initiatives can make a difference in that. Right from the beginning, when researchers are starting out on their research path, having that training and support on how to do things in an open way can prevent problems down the line. I really like a paper written by Florian Markowetz in which he talked about five selfish reasons to work reproducibly to protect yourself, to co-ordinate your data and not to lose it, and to do things right, which can set you up for all the other steps that follow.

Q124 **Katherine Fletcher:** Dr Dhand, do you want to add anything on what the industry needs to do—things that we should look at further?

Dr Dhand: We need to mandate being open, which is what everyone has been saying. We have standing bodies that have mandated open access. That is just the paper. They should be mandating open data deposition and that the protocols are written on every single paper and deposited in a repository or in a journal that publishes protocols. Publishers have journals that will publish protocols, or we have repositories for open protocols. The protocol is absolutely the foundation for reproducibility. Open data does not require just data deposition; it requires detailed notes on how that dataset was created so that it can be used—

Q125 **Katherine Fletcher:** Which should exist in a quality research environment.

Dr Dhand: We have journals that will publish these descriptive papers on datasets, but not all of the community are on board.

Katherine Fletcher: That is incredibly helpful. I appreciate it. I am going to have to pass back to the Chair.

Chair: Thank you very much indeed, Katherine. We need to wrap up this session. Can I thank Dr Dhand and Dr Moylan for very comprehensive answers on the role of publishing in these questions? It is very helpful, as we write, our report to draw on your evidence today. Thank you for helping the Committee with its inquiry.

Examination of witnesses

Witnesses: Richard Horton, Viscount Ridley and Dr Chan.



Q126 **Chair:** We turn to our next and final panel of witnesses this morning. For this session we are looking at a particular ongoing research question, if I can put it that way, that entails matters of discovery, investigation and publication, and that is the origin of Covid-19.

To help us to begin to ask some questions on this, we have before us Richard Horton, editor-in-chief of *The Lancet*. Mr Horton, welcome back to the Committee. We also have with us Lord Ridley—Matt Ridley—the author of a book entitled “Viral: The Search for the Origin of Covid-19”, which was recently published, and his co-author, Dr Alina Chan. I believe Dr Chan is dialling in from Boston, where it is about six o’clock in the morning. Thank you for giving evidence at such an early hour.

Perhaps I can start with a question to Richard Horton. On 18 February 2020, crucially, *The Lancet* published a letter that denounced “conspiracy theories suggesting that Covid-19 does not have a natural origin”. There were various follow-ups to that, which we will come to. To go back to the first publication of that letter, could you talk us through how it came about and what was the process behind it? How did it come to your attention and what were the discussions about publishing it?

Richard Horton: Thank you for inviting me back. On that particular letter, it is worth reiterating the title. You have rightly pointed to one part of the letter, but the main reason for publishing it lies in the title, which is “Statements in support of the scientists, public health professionals and medical professionals of China combating Covid-19”.

There were two main reasons why it came to my attention. First, there was the political context. In the period late January through early February there was growing criticism of China by politicians, particularly in the United States. For example, Tom Cotton, a Republican, denounced China for being duplicitous and dishonest in its reporting of the pandemic. There were initial claims being made about the possibility of an escape of the virus from the Wuhan Institute of Virology.

The second reason was a more public context, in that there was growing anti-Asian sentiment. The Asian American Journalists Association reported literally thousands of xenophobic or racist attacks against Asians. In many European countries, including the United Kingdom, there were physical attacks against people of Asian descent because of the narrative around the origin of the pandemic.

Therefore, the purpose of publishing this letter was to say that, instead of blaming Chinese scientists or China generally, we should be offering our solidarity—I think that was the word used in the letter—with colleagues in China to try to get to the bottom of what caused the pandemic.

The second part is what you say: be careful about raising speculations at this early stage because we do not have any evidence one way or the other about the possibility of a lab leak.

Q127 **Chair:** This was very early in the pandemic; it was 18 February 2020.



HOUSE OF COMMONS

How much of this was justified and how much interest was paid at the time, given that this was a letter rather than a refereed journal article, to the authors' potential conflicts of interest?

Richard Horton: We ask everybody who submits a piece that is accepted for publication in *The Lancet* to declare their competing interests. We take those statements on trust, and in this particular case regrettably the authors claimed they had no competing interests. Of course, the implication of your question, as you well know, is that there were indeed significant competing interests, particularly in relation to Peter Daszak, who was the leader of the EcoHealth Alliance.

Q128 **Chair:** As of the 18 February publication you were not aware of those competing interests.

Richard Horton: We were not aware of those competing interests, but we very quickly became aware of them because he was subject to considerable public criticism for signing this letter and saying that there were no competing interests. We ended up having a debate with him about whether he did or did not have a competing interest. It was quite an interesting debate to have. His view was, "I am an expert in working in China on bat coronaviruses. That is not a competing interest. It actually makes me an expert with a view that should be listened to." Our take was, "That might be your view, but in the court of public opinion that is a competing interest you should declare." It took us over a year to persuade him to declare his full competing interests, which we eventually did in June of this year.

Q129 **Chair:** You have explained that you have a self-certification or self-declaration model.

Richard Horton: Yes.

Q130 **Chair:** Would it nevertheless have been reasonable for you either to have expected to know about his prospective competing interests or inquired yourselves and looked into that?

Richard Horton: If we had infinite staff then perhaps, but we are publishing an issue every week of about 90 pages with literally hundreds of authors. We do not investigate every single author that we publish for their possible competing interests. The whole process of scientific publication—this relates to some of the issues you have been discussing in the previous two sessions—depends, rightly or wrongly, justly or unjustly, on an element of trust. We trust authors to be honest with us, and authors trust us to deal with their work confidentially and appropriately. Sometimes that system breaks down, and in this particular case Peter Daszak should certainly have declared his competing interests right at the beginning.

Q131 **Chair:** On 21 June 2021 you published an addendum to that letter of February.

Richard Horton: That is right.



Q132 **Chair:** What precipitated the publication of that addendum?

Richard Horton: As I mentioned, we were engaged in a series of discussions with him over the previous year about what constituted a competing interest. There was a difference of opinion between him and us about that subject, but eventually I think he recognised that the debate in the court of scientific and also public opinion was such that explaining the nature of his relationships with the EcoHealth Alliance and the work it was doing in China was of material interest and of extremely important relevance to interpreting his letter of 18 February.

Q133 **Chair:** Would I be correct in interpreting your response—I do not want to put words into your mouth—that, with hindsight, you think that information should have been included in the original letter of 18 February?

Richard Horton: I completely agree, 100%. The information we published in June as an addendum should definitely have been included in the 18 February letter. You are absolutely right.

Q134 **Chair:** As you have heard, our investigation is into research methods, integrity and publication generally. Have you considered making any changes to the publication process to guard against something like this arising in the future?

Richard Horton: On the specific issue of competing interests?

Chair: Yes.

Richard Horton: We ask authors of every piece of work that is submitted to declare their competing interests. Have we changed our methods there? As I say, we are publishing literally hundreds of authors every week and we do not investigate every single author ourselves; we take on trust what they say, but certainly our awareness has been heightened by this issue. Certainly, in the context of Covid, we are now more vigilant about who we are publishing and why we are publishing, and what their potential conflicts might be.

Q135 **Chair:** On 17 September of this year *The Lancet* published a letter that included an appeal for “objective, open, and transparent scientific debate about the origin of SARS-Cov-2”. Was this the result or reflection of a change in editorial policy by *The Lancet*?

Richard Horton: No, I do not think so. We published several letters in September relating to this, including for the first time a letter from Chinese authors, which was really quite revealing. What happened was the WHO investigation, which I think was important. In February there was a lot of speculation about the validity or otherwise of the laboratory leak hypothesis. The WHO in May passed a resolution to establish an independent investigatory team to go to Wuhan in China to try to understand the possibilities for the origin of the pandemic. It took most of 2020 to put that team together and to get it accepted by the Chinese authorities.



It went to China in the early part of 2021 and published its report in March. In its report it identified four possible pathways for how the pandemic could begin. For the very first time, the laboratory leak was officially endorsed by the WHO as a possible pathway for how the virus got into the human population. After that moment in March, when it got that official stamp of approval as a valid hypothesis, it opened the door for a much more transparent debate about what the scientific evidence was for and against a laboratory leak. I believe that until the WHO report in March it was still highly speculative, but after the visit took place, during which the Wuhan laboratories were assessed and interviews took place with Chinese officials and scientists, there was a step change in the debate. Now we have phase 2 of the WHO investigation, which will begin very soon to try to explore one level deeper those four pathways.

Chair: Thank you very much indeed, Mr Horton.

Q136 **Aaron Bell:** Mr Horton, thank you for coming to us again. I remember you coming to us early in the pandemic. You have obviously taken a strong interest in it. For the record, you said in answer to the Chair that when you published the letter on 18 February you had no knowledge of Dr Daszak's links to the Wuhan Institute of Virology.

Richard Horton: That is absolutely correct.

Q137 **Aaron Bell:** You referred in your opening remarks to xenophobia, anti-Chinese attacks and so on. We all deplore that sort of thing, but it is not the function of a scientific journal to combat that. The function of a scientific journal is presumably to illuminate the truth and try to get to the true science behind things. Do you think that the publication of a letter served those ends, or did it serve to close down the scientific debate prematurely?

Richard Horton: There are two elements to your question and perhaps I might deal with each of them. First, the goal of the letter primarily was to say that we have a global pandemic and the solution to it lies in global co-operation. That means that we should see the Chinese medical and scientific community as partners in this endeavour rather than blaming them. What was taking place in those early weeks and months was a bit of a blame game, which of course is still going on to some extent. The purpose of the letter was to say, "Let's make common cause with our Chinese colleagues to try to get to the bottom of the origins of the pandemic."

As to the silencing effect, I am certainly aware that that has been raised, but the solution to understanding the outbreak in Wuhan lay in studying what took place in Wuhan. The World Health Organisation in the early months of 2020 very quickly put together the resolution that kick-started the process of its independent investigation. There was no slowing down of that process; there was no silencing of the WHO; there was no silencing of the investigation team that went to Wuhan.

Q138 **Aaron Bell:** Mr Horton, the WHO team that went to Wuhan was not



HOUSE OF COMMONS

allowed full access and produced a report that was clearly informed by as much as the Chinese allowed them to see. I applaud the idea of co-operating with China, but it is clear that it has not been open and straight with both the WHO and the wider scientific community throughout, is it not?

Richard Horton: You are absolutely right. It was denied access to raw data that it deemed materially crucial to its investigation to elucidate which of the four possibilities was most likely to be true. What it did—this is not unimportant and it wrote about it recently in *Nature*—was identify the laboratory leak hypothesis as a perfectly legitimate and valid hypothesis that needed to be tested and understood. Until that point, that idea had been denied by the Chinese Government and many others. That WHO team put it on the table and changed the terms of the debate.

Q139 **Aaron Bell:** Mr Horton, with respect, and I will come to our other witnesses, the fact that the virus originated in Wuhan, where the Wuhan Institute of Virology is located, makes a prima facie Bayesian case that there was a possibility of a lab leak in the first place. Our other witnesses who have written this book said that even at the beginning they did not necessarily believe it was a lab leak but just something that needed to be considered. Did you not believe that it should even have been considered in January/February 2020? Have you had to wait for the WHO to say that it should be considered for you to take that position?

Richard Horton: What we had in the early part of 2020 was nothing but speculation about the possibility of a laboratory leak based on the background understanding of 20 years that many bat-related viruses had come through zoonotic infections. We published several commentaries in the early part of 2020 discussing the likelihood of a zoonotic infection.

Q140 **Aaron Bell:** The theory was based on a wet market, not bats.

Richard Horton: Yes, but it was proposed that the intermediate host was one of the animals that could have been in the wet market. If you look at Michael Worobey's work—Dr Chan signed a letter with David Relman and others calling for more transparent investigations earlier this year in *Science*—his conclusions, published most recently in *Science*, are that the preponderance of cases in that market suggests that a critical event took place there that would provide evidence to support a zoonosis, not a laboratory leak. I absolutely agree with you that these questions still remain open and demand further investigation.

Q141 **Aaron Bell:** You say they remain open, so clearly you think that a lab leak is now a credible hypothesis. Would you go as far as to say that it is the likely hypothesis? What probability would you attach to that based on all the knowledge you have accumulated over this pandemic?

Richard Horton: I would agree with the WHO conclusion that it is a hypothesis that should be taken seriously and needs to be further investigated, but it deems that hypothesis as extremely unlikely compared with the possibility of a zoonotic infection through an



HOUSE OF COMMONS

intermediate host, which it deems much more likely. I think that was an entirely justifiable conclusion. The evidence we have had since then, at least to my reading, supports the view that it expressed in March.

Q142 **Aaron Bell:** Would you care to put a percentage on the probability of a lab leak?

Richard Horton: No. I would be happy to agree with the WHO's view that it is extremely unlikely in comparison with a zoonotic infection.

Q143 **Aaron Bell:** Perhaps I may turn to the two other witnesses, Lord Ridley and Dr Chan. Thank you for appearing before us.

Dr Chan, what probability would you put on the possibility of a lab leak as the origin of Covid-19?

Dr Chan: I cannot give a number. I think a lab origin is more likely than a natural origin at this point. We all agree that there was a critical event at the Wuhan seafood market and that it was a super-spreader event caused by humans. There is no evidence pointing to a natural animal origin of the virus at that market.

Before I get into that, I would like to point out that *The Lancet* really needs to publish all of the manuscripts it received originally, whether they were withdrawn or rejected by the journal. We need to see what it received from Chinese scientists in the early days of the pandemic. We know from Jeremy Farrar's book that *The Lancet* was in possession of information on human-to-human transmission of this virus and symptomatic transmission of it. It did not release that to the world. Many more lives could have been saved if it had been released to the world. I came prepared with five recommendations for journals when dealing with issues where withholding or sharing data can be measured in the numbers of lives lost or saved.

Chair: Dr Chan, I am sure we can come to them, but through questions from Members.

Q144 **Aaron Bell:** I will put to Mr Horton later your suggestion about the manuscripts. You say it is more likely than not that it was a lab leak. How confident are you that we will be able definitively to determine the origins of Covid-19 over time?

Dr Chan: I am very confident. We have seen from previous cover-ups that it just takes time. Right now it is not safe for people who know about the origins of this pandemic to come forward. It might be five years or 50 years from now, but we live in an era when so much data is being collected and sought. Everyone has a phone with a camera. Email messages were flying out of Wuhan during the early days of the pandemic. We need a credible system of investigation to collect all of these pieces of evidence and put them together to get a much better understanding of how this pandemic might have started.

Q145 **Aaron Bell:** I will turn to your co-author. Lord Ridley, what percentage



chance would you put on Covid having originated as a lab leak?

Viscount Ridley: Like Dr Chan, I do not like putting a number on it, but I also think that it is now more likely than not. We have to face the fact that after two months we knew the origin of SARS through markets. After a couple of months we knew the origin of MERS through camels. In this case, after two years, we still have not found a single infected animal that could be the progenitor of this pandemic. That is extremely surprising. Eighty thousand animals have been tested all across China and none was found in that market. That was announced as early as May 2020 and it was one of the things that got me intrigued by this possibility.

We know of one animal that brought a closely related virus from a bat cave in southern Yunnan to, specifically, the city of Wuhan in the years before this. That animal is homo sapiens in the shape of scientists. Scientists were travelling 1,000 miles or more to Yunnan to collect SARS-like viruses and bring them back for experiments in Wuhan. To me, it has to be taken seriously. It is regrettable that in 2020 there was a pretty systematic attempt to shut down this topic.

Mr Horton says that the main purpose of that letter was to combat anti-Chinese conversations. That can be done while discussing the origin of the pandemic. After all, whether or not it came out of a market or laboratory in China should not affect the position of China as the source of this pandemic.

Q146 **Aaron Bell:** I assume both of you think that it was an accident that it leaked from a lab. You do; both of you are nodding. Within that, Lord Ridley or Dr Chan, do you think it is more or less likely that the virus was modified in the lab before it escaped potentially through gain-of-function research?

Dr Chan: We have heard from many top virologists that a genetically engineered origin of this virus is reasonable; they have said it is worth investigating. This includes virologists who themselves made some genetic modifications in the first SARS virus. We know now that this virus has a unique feature, called the furin cleavage site, that makes it the pandemic pathogen that it is. Without that feature, there is no way this virus would be causing the pandemic.

Only recently, in September, a proposal was leaked showing that scientists in the USA in the EcoHealth Alliance were in collaboration with the Wuhan Institute of Virology developing this pipeline for inserting novel furin cleavage sites—these genetic modifications—into novel SARS-like viruses in the lab.

As to what we have right now, the analogy I use is that these scientists said in early 2018, “We’re going to put horns on horses,” and at the end of 2019 a unicorn showed up in their city. It is a striking coincidence that needs to be investigated. I would say that the burden is on scientists to show that their work did not result in the reaction of SARS-CoV-2.



Q147 **Aaron Bell:** Dr Chan, do you think it is more likely than not that the virus was modified in the lab before it leaked?

Dr Chan: I am saying it has to be investigated and we can investigate it because the EcoHealth Alliance has these communications and documents with the Wuhan Institute of Virology that can tell us their thinking and the experiments they were considering. How did they come to write this proposal in 2018 saying they were going to insert novel furin cleavage sites into novel SARS-like viruses in the lab?

Q148 **Aaron Bell:** Lord Ridley, you wanted to come in there.

Viscount Ridley: This goes back to something you discussed in an earlier session, namely that the grant proposal to DARPA to do this work had to be leaked. We did not know about it otherwise. It is pretty extraordinary that the information about the plan to do these experiments, which are absolutely critical to this question, was not revealed by Dr Daszak, who was the principal investigator on the grant application, until it was leaked a couple of months ago.

To go back to the questions about Dr Daszak's role in *The Lancet* letter, not only was he one of the co-authors; he orchestrated it. Again, we had to find that out from leaked emails. He said to his co-authors that it would not appear to be coming from him or his organisation, and he remained on *The Lancet* commission investigating the origin of Covid for many months thereafter. So there has been a significant lack of transparency not just from the Chinese authorities but from western ones as well in this, and that does seem to me to be a huge problem in the context of your inquiry into the importance of scientific transparency.

Q149 **Aaron Bell:** What would you describe as the benefits of determining the origins of Covid-19? If it does prove to be a lab leak, what are the potential consequences for both science, including western science, and international relations?

Viscount Ridley: I think we need to find out so that we can prevent the next pandemic. We need to know whether or not we should be tightening up work in laboratories, or whether we should be tightening up regulations relating to wildlife sales in markets. At the moment we are really not doing either. We also need to deter bad actors who are watching this episode and thinking that unleashing a pandemic is something they could get away with and they would get pretty much a Potemkin report from the World Health Organisation if it happened.

As for the impact of finding out, if we did find out this was a laboratory leak, obviously it would have huge implications. It is important to distinguish the enormous benefits that we get out of biotechnology research and how much enormous good has come from that, including the vaccines that are helping us to survive the pandemic, from the fact that one or two experiments seem to have been done. Whether or not this was the cause, we now know that experiments were being done at biosecurity level 2 in Wuhan that resulted in up to a tenthousandfold



increase in infectivity in viruses, or a three or fourfold increase in lethality in humanised mice. These are the kinds of experiments that will play into the hands of those who are critical of science and want to stop this kind of research. The important thing is to stop doing experiments of this kind that are risky while continuing to do ones that are less risky. For that, we need much greater transparency across the world about what kinds of gain-of-function experiments are being done on viruses.

Q150 Aaron Bell: Perhaps I may return to Mr Horton to give him right of reply to some of the things other witnesses have said. Dr Chan challenged you to publish all the manuscripts you received, including those from Chinese institutions, over the course of the pandemic. Is *The Lancet* willing to do that, Mr Horton?

Richard Horton: I am not quite sure what that would achieve. The usual practice is that the communications we have with authors are confidential, so I do not see the value of going back to all of those authors and seeking their permission to disclose that they submitted work to us. If we are talking about the context of the February letter, we dealt with that directly. If we are talking about the hypothesis of a lab leak raised by Dr Chan and Lord Ridley, that is an entirely reasonable hypothesis that needs to be further investigated. I think that the WHO team that has now been put together to go to Wuhan to continue those investigations is where we should be looking for answers.

Q151 Aaron Bell: Lord Ridley said that Peter Daszak remained on your *Lancet* commission on coronavirus for a very long period of time after you were presumably aware that he was extremely conflicted in writing that letter in the first place. Why was he allowed to remain on your commission?

Richard Horton: You are absolutely right; he did remain on the commission. We established the commission in the summer of 2020 and it began its work towards the end of 2020. It is led by Jeff Sachs based at Columbia. When Peter Daszak's conflicts of interest became known to us and became a concern to us as we found out more, we raised it with him and the taskforce we had put together on Covid origins.

Q152 Aaron Bell: You said it became known to you very quickly after the letter because it was immediately the source of media interest.

Richard Horton: That is true, but we were not aware of the extent of that conflict until we had done our own investigations. There had been a lot of discussion in social media and in the press about it, but we needed to know the details of exactly what his alleged conflict was. We were trying to get that information from him and that proved to be difficult for several months. When we eventually did get that information and raised it with him and the taskforce, it became clear that that taskforce was not going to be able to pursue an independent inquiry into origins and we disestablished it.

Q153 Aaron Bell: That was why you disestablished the taskforce. It took 16 months for you to publish that addendum. Dr Chan has described that



HOUSE OF COMMONS

as too little too late. I have to say I agree with her. Do you not think it was too little too late?

Richard Horton: I think I explained that in answer to the Chair's earlier question. When we went to Dr Daszak to ask about this competing interest we ended up having a dispute.

Q154 **Aaron Bell:** With respect, Dr Horton, we are in a fast-moving pandemic where trust in science is crucial. You cannot take 16 months to resolve a basic conflict of interest that was apparent from the moment the letter was published.

Richard Horton: I completely agree that we needed to move fast. We moved as fast as we could, given that this particular individual disagreed with us about the nature of his conflict of interest. We needed to get this information accurately described and for him to agree a text that described his relationships with the Wuhan Institute of Virology, the EcoHealth Alliance and the research on bat coronaviruses. We eventually extracted that and published it in June, as you have already described. I think we did a good job of correcting the record as quickly as we could.

Aaron Bell: When you appeared here before, Mr Horton, you expressed the need for things to move quickly in a pandemic. This seems to have taken far too long. I hope that Dr Daszak will be willing to come and give evidence to this Committee in future so that he can explain for himself why he delayed your inquiry into his conflict of interest for so long; it took 16 months. Anyway, I will leave it there and return to the Chair.

Q155 **Graham Stringer:** Thanks for coming this morning. Lord Ridley, I have read most of the reviews of the book written by you and Dr Chan. I have not finished the book yet, but I have read parts of it. The reviews accuse you of stretching facts and sensationalism. Would you like to respond to that?

Viscount Ridley: We think that is a very odd characterisation of our book. I hasten to point out that we have also had good reviews.

Q156 **Graham Stringer:** I was not going to ask you to comment on the good reviews.

Viscount Ridley: We were very careful in this book never to put anything in that we could not evidence to some degree. Every now and then I would write a paragraph of speculation, saying that maybe they were doing this or that or something. Alina would be very strict and say, "Sorry. That paragraph has to come out because we have no evidence for it." About one fifth of the book consists of references. We back up every statement we make. We give both hypotheses equal time in the book, and at the end there is a chapter where we hand the microphone, as it were, to a lawyer to make the case to the jury that it began in a market. When I reread that chapter I find it quite persuasive. We then hand the microphone to a lawyer to make the opposite case, and again I find that quite persuasive. So we think we are very fair on both hypotheses. The



HOUSE OF COMMONS

one thing we do not do is go into speculation about bioweapons or the political aspects of things.

To give a specific example, there is a report from inside the US intelligence community that three of the first cases were workers at the Wuhan Institute of Virology who got sick in November 2019 with symptoms very like Covid-19. There are some very specific details attached to that. We mention that but say, quite clearly, that because we do not have security level clearance we cannot check whether or not that is true. We are very clear about what we can and cannot confirm, and the book is rigorously factual and the very opposite of sensationalist.

Q157 **Graham Stringer:** If the scientific debate had been operating, as I think all of us would wish, in an open, transparent, quick and speedy way, would there have been any need to write this book?

Viscount Ridley: Frankly, no. The motivation to write it driving Dr Chan and me was the fact that it was very clear that we did not know the answer to where this came from, and that was because of a lot of efforts to obscure details. The changing of the names of viruses, providing no reference to where it was found and things like that right at the beginning meant it took several months to uncover the source of RaTG13, the most closely related virus at that time to SARS-CoV-2. All of these obscurings were a red rag to our bull. We wanted to try to find out what was going on. We would much rather have been in a situation where it had become very clear very early on and people had been invited in to investigate all of the details and rule out the laboratory and the market or whatever.

There is a database at the Wuhan Institute of Virology with 22,000 entries in it, 15,000 of them relating to viruses from bats. It has been offline since before the pandemic. It was there to help prepare for pandemics. Which pandemic are they waiting for before they share it with the rest of the world? Quite a lot of the entries in it relate to viruses collected with US Government funding through the EcoHealth Alliance. Why do they not have access to that data, which goes back to the earlier session about the importance of making data available?

Q158 **Graham Stringer:** Dr Chan, I have read that co-authoring this book has led to some personal difficulties for you and threats. Is that the case? If it is, can you tell us what has happened?

Dr Chan: It is true that that is the case, but I do not think it is good to get into details about this because it distracts from the matter at hand. I would rather spend the remaining time talking about why it is important to get those original manuscripts. This is very important to me, and it should be the first priority of an investigation into the origin of Covid-19. We know that lots of scientists inside China were sending out information in the form of manuscripts in the early days of this pandemic, not just to *The Lancet* but to many other prominent, prestigious journals, some of whose editors are here today. We need to see those. We need to know what scientists knew and what they were sharing at the beginning before



HOUSE OF COMMONS

the gag order came in and they started withdrawing their papers and altering their papers. This will really help us to fill out the picture of what was happening.

Q159 Graham Stringer: I appreciate that, but part of science working is that scientists have to have freedom of thought or freedom of movement. It seems to me from what I have read that there are attempts to restrict your freedoms. Is that the case?

Dr Chan: I would say that is the case for a lot of scientists who are handling Covid-19 issues. This is so controversial that anything—masks, vaccines, whether the virus is airborne, even very basic things like that, not to mention the origins of Covid-19—will result in threats against scientists, but it is unavoidable. I am not saying it is right, but I am not in a rare situation. A lot of scientists have suffered a lot of abuse, but in this situation specifically there are potential career effects. For a scientist to come out and say something that the rest of the community does not want to talk about, for it to be condemned as a conspiracy theory since early 2020, and for anyone raising the possibility of a lab leak as the origin to be regarded as anti-scientific, racist or a right-winger, is crazy. This is a scientific problem and it cannot become policy that we can investigate only lab-based outbreaks in white countries.

Q160 Graham Stringer: Mr Horton, I was surprised by your answers to my colleague Aaron Bell. You said that, once the World Health Organisation had said the lab leak theory was a possibility, that somehow validated that hypothesis. Why do we need the WHO, which is a highly political organisation as well as a health organisation, to validate a hypothesis?

Richard Horton: It was the result of the first independent investigation into what took place in Wuhan. For all its imperfections and its inability to get at raw material that it said it needs further access to, it was the first time that an official independent investigation had put the laboratory leak hypothesis on the table. Up to that point there had been a lot of debate and speculation about this, but I think that was a turning point in placing that hypothesis as a serious contender with three others for further investigation.

I understand the charge that Chinese scientists might have been gagged in some of their discussions of this, but one of the letters that we published in September of this year was from a group of Chinese scientists in fact, including the president of the Chinese Academy of Medical Sciences, and that letter argued that the laboratory leak hypothesis was extremely unlikely but it does not 100% rule it out. It says there needs to be international co-operation to properly understand the origins of the pandemic. Within the political constraints that Chinese scientists live with, there is an acknowledgment that there is an area of uncertainty here that needs to be investigated, and that comes from some of the highest echelons of Chinese science. That is why the WHO phase 2 investigation becomes so important.



Q161 **Graham Stringer:** Again in answer to questions that Aaron asked, you said that you had to trust your contributors on conflicts of interest. Was nothing learned about trust in *The Lancet* from the experience of Wakefield and the MMR autism hypothesis, when it took 12 years to withdraw a fraudulent paper?

Richard Horton: One of the lessons there was that you need to have independent scrutiny when allegations of misconduct are made. There needs to be proper due process where those investigations can take place. That is one of the lessons that we drew from that whole case that you mention. I am not sure how relevant it is to the Peter Daszak competing interest issue here.

Q162 **Graham Stringer:** It is relevant to how you assess papers or letters that are put forward to *The Lancet*. You seem to be saying that we have to trust what is put forward. From the beginning, partly because it is China with its political system, this has been an area of controversy, so just to accept it on the basis of trust, particularly given the trust you put into the Wakefield paper, requires some justification, I think.

Richard Horton: Your question is reasonable given that Peter Daszak had those connections with the EcoHealth Alliance and the Wuhan Institute of Virology, but, as I say, the edifice of science, rightly or wrongly, depends to a large degree on trust. When papers are submitted to us for peer review, we take it on trust that the description of research that has been presented to us is an accurate description of what took place. We do not go back to that institution and check the raw data. We do not look at the primary records. We do not look at the case reports of randomised trials to make sure that all of those data are indeed true. We take it on trust that that information is correct.

If you are arguing that that is a step too far and that we should be investigating, that requires a radical change in our publication system. I hope that we have enough confidence overall in the institutions of science that we can agree that episodes of outright research fraud are relatively uncommon and do not require such a regulatory bureaucracy that would impede science and the speed with which science delivers important answers.

Dr Chan: I would like to jump in. I think there is too much trust and this trust is being exploited sometimes by bad actors.

Q163 **Chair:** I am sorry, I missed that, Dr Chan. Would you repeat what you said?

Dr Chan: It is great that journals are so trusting, but sometimes in these times of crisis this trust is being exploited by bad actors to shape the narrative and to shut down discourse. They send out fake data to mislead people into thinking this does not spread among humans. This is crazy. We cannot trust everything in times of crisis. I am going to say that about the WHO as well. Let us be clear about how it was decided whether the bat leak was likely or not. They went into a room where there were



Chinese Government officials and asked the scientists at the Wuhan Institute of Virology, “Did you do this?”, and they said, “No, we didn’t.” Then they voted to see how likely it was, in front of Chinese Government officials. What did they think the response would be? Who would say, “We think a lab leak is likely”? Let us be clear that this was not a scientific process.

Chair: Our inquiry is on the publication practice of journals and the way that scientific research is conducted. In this respect *The Lancet* of course is an important scientific journal, so we have some questions relating to its role as a scientific journal following on from the questions asked of our earlier witnesses. Rebecca Long Bailey is going to ask some of these.

Q164 **Rebecca Long Bailey:** Thank you all for speaking to us today. My questions are to Richard. What role is *The Lancet* currently playing in ensuring transparency in research?

Richard Horton: I think what we are trying to do in terms of broader issues of supporting research integrity, of which transparency is a very important part, extends from, for example, the peer review process in which we ask reviewers to review not just the paper—the five-page essay that you heard about earlier—but also the protocol that goes with the paper, which is of course a considerably larger document, together with the statistical analysis plan, so that there is a much fuller understanding of the nature of the study.

We, of course, have moved into the area of pre-prints and we have our own pre-print server that tries to encourage that early publication of work. That has been extremely important during the pandemic so that journals are not justifiably accused of sitting on work for a long period before it is in the public domain, and people can have access to an early version of it. That improves transparency.

The whole open access movement has been extremely important in improving transparency, and certainly at *The Lancet* over the last few years we have launched a range of open access journals to try to improve transparency and access to information.

On the broader issue, and this was raised by Ben earlier, the methods of science have, to some degree, outstripped the way journals work. In the old days, going back 30 years or so, you could get published in a journal with a 2x2 table, very simple epidemiology. It does not work in the same way now. You often have very complex models and computational statistics that are actually very difficult to reproduce. That is where journals struggle to make that information transparent and where we have to work with institutions so that they provide that information independently of journals.

In terms of rigour for us, at a clinical journal, we are trying to publish studies particularly in the clinical trial arena that do not suffer from high levels of random error. Large simple trials can reduce the risk of



mistakes. Again, if you go back 20 years or so, many of the early randomised trials were often small, single-centre studies that raised many false positives, and we have tried to learn lessons from that.

Q165 Rebecca Long Bailey: That is really helpful. How easily can the data underpinning a research paper be made available on your platform? There was a problem raised earlier about how peer review never takes place on the software that is used to collate a piece of research. How would you propose trying to overcome that in a journal?

Richard Horton: This is a problem we are struggling with right now. I do not have a simple answer that I can give you, because it is a current challenge. Let me give you a very live example at the moment with Omicron. Last Friday, I was discussing with a particular modelling group in the United States how we could model and forecast the pandemic globally factoring in Omicron. The answer to that question is that you have to factor in infection-related immunity, vaccine-mediated immunity, waning immunity and the severity of the disease. In several of those areas we simply have no data. I asked the person I was talking to, "What are the mechanics of the model that you have? How are you putting the model together?"

The complexity is just enormous because you have 194 countries. You are trying to get data from those countries on the extent of the pandemic—often sub-national data. You are trying to put that into a computational model with the four variables that I have just mentioned, which are overlaid on Delta, which has been the previous dominant variant. The complexity is enormous. How do you make that transparent and reproducible for another independent research group? I do not have an answer to that question because those data are not fully available for everybody to see. I think all modelling groups are struggling with this at the moment: making those data available for independent scrutiny and independent validation. It is a struggle. This is the point that I made just a moment ago that, in a sense, the science—particularly the modelling science—is outstripping some of the mechanisms that we have for independent validation and transparency.

Q166 Rebecca Long Bailey: Thank you. That is really helpful. Dr Chan.

Dr Chan: May I quickly jump in? Scientific journals can do a lot to help research integrity. I do not know why everyone here is running out of ideas. I have some really good ones for you.

Make pre-prints mandatory prior to submission; in that way everyone can see it. Check that all data is deposited on international databases that cannot be altered by the submitters later; we have seen this happen where the Covid-19 sequences were deleted off the international database. Journals can do this; they can just check it. Publish peer reviews; you can keep it anonymous but make it open so that people can see if there is gatekeeping or bias. Refuse to publish papers when novel pathogen sequences have not been shared in databases for more than a



year after discovery. Do not publish any more work where they are hiding pathogen sequences for years and years. Incentivise replication studies and timely publication of submissions that correctly call out errors in papers. These steps will immediately help the research community see quickly when there are errors in papers, when data is missing and when there are incorrect sample histories.

Rebecca Long Bailey: Thank you, Dr Chan.

Q167 **Chair:** Perhaps, Mr Horton, you might give a response to some of the suggestions that Dr Chan has given.

Richard Horton: I think they are extremely constructive suggestions. In the world that I work in, which is medicine and global health, the issue about making all data available immediately on publication can be a little challenging because oftentimes data are provided by Governments under confidentiality clauses to researchers, and it is not within the power of researchers to make those data publicly available. We have run into problems with that in the past where we have had a request that an organisation provides data for independent scrutiny by another research group and the authors will tell us that they are contractually unable to share that information.

I think it is an aspiration—an ideal—that we should certainly be working to, but the reality, unfortunately, of doing international collaborations is that there are sometimes contractual details that prevent that.

Viscount Ridley: Is there time for a very quick comment?

Chair: Of course, Lord Ridley.

Viscount Ridley: To give an example of the system working well in this case, a paper was published, which was submitted before the pandemic began but published after that, by the EcoHealth Alliance, Latinne et al. It was a summary of 630 coronaviruses that had been studied up to 2015 by the group. What was interesting was that they uploaded all the partial genome sequences into an international genetic database. This enabled two citizen scientists—Francisco de Ribera and an anonymous person called BabarElephant—to identify in June 2020 that there were eight viruses very closely related to SARS-CoV-2 within the sample that had been collected from the same site in Mojiang where the most closely related one had come from and that were being held at the Wuhan Institute of Virology. It took another six months before the existence of those eight viruses was confirmed by the Wuhan Institute of Virology in its addendum to the *Nature* paper. Do not neglect the importance of individual citizen scientists going out and doing the dirty work of finding out what is in these databases.

Q168 **Chair:** Thank you, Lord Ridley, that is a very good point. This Committee is followed by many citizen scientists in the country and no doubt they will have heard that call to arms, which I am sure the Committee will



endorse.

Finally from me on this point of publication before I turn to my colleague Katherine Fletcher, I do not know whether you heard Dr Moylan in the previous session, Mr Horton, refer to publication practices. She made a remark that I am surprised by, that journals are as likely to publish confirmatory studies as breaking news, as it were. Does that apply to *The Lancet* and its stable of publications?

Richard Horton: Yes, I heard that. Let us take a very specific example. We had the opportunity to publish the very first AstraZeneca vaccine trial, I think, in December 2020. If we then received the fifth randomised trial looking at the AstraZeneca vaccine, would we have published that in *The Lancet*? The answer honestly is that it is very unlikely.

What we are asking is: is the paper asking and trying to answer an important scientific or clinical question? That means that it could be the first study, like the AstraZeneca paper, or let us say there have been several papers on AstraZeneca and they have come to slightly different conclusions or raised other questions. Perhaps we would have published the fifth trial if there was an important scientific or clinical question, but for a general medical journal, of course, we are making judgments all the time about what is important and what is not important. To say that we are publishing absolutely everything that is sound science, clearly we are not. We have to make choices.

On the front of *The Lancet*—we still have a print copy, amazingly—it says that we are a newspaper, and we have been a newspaper for 200 years. We are just a very specialised newspaper. Just like newspapers have to choose what they think are the most important stories for their readership, so we are making judgments about what we think are the most important areas of science for our readers. That requires judgment. I hope that we get it right most of the time. Sometimes we get it wrong, but that is what a journal does. We are not an electronic database that publishes everything. We are a journal with a specific community that we are trying to serve, so we do make judgments and we do make choices.

Q169 **Chair:** That conforms to my expectation. You have to attract interest to your journal, which is why I was surprised by that piece of evidence. But do you recognise the structural problem that some of our earlier witnesses pointed to that it may be socially beneficial, if I can put it that way, for people literally just to reproduce and validate conclusions that someone else has drawn, and if the research design and integrity behind it has all been proper, as we would hope, actually you would not find anything terribly interesting per se and that will put people off, especially early-career researchers who want to have some articles in prestigious journals? There is a structural problem here, is there not?

Richard Horton: There is a problem. I can only speak for *The Lancet* group of journals. We have over 20 journals at *The Lancet*. Let me take my AstraZeneca example again. The first trial comes to *The Lancet* and



we publish it. Let us say there are subsequent trials. They will be very interesting perhaps not to a general audience but to a specialist audience. We have an infectious disease version of *The Lancet* and we ask the authors if they would consider publishing in *The Lancet Infectious Diseases*. If they say yes, it gets published there, subject to peer review.

We try to transfer papers across *The Lancet* journals so that they are appropriate for a particular audience. The peer review is not wasted. The peer reviews we might have done for the weekly are then used for our other journals. In that way, we can have an efficient means of not falling into the trap that you rightly describe. There is an enormous amount of waste in the system, and what publishers have to do is figure out how to reduce that waste. We have one solution and it works pretty well for us, but I absolutely acknowledge the challenge that you are drawing attention to.

Chair: Thank you very much. Katherine Fletcher, finally, in this session.

Q170 **Katherine Fletcher:** This inquiry is about the integrity of research and what we can do to make sure that we have the gold standard. I have been listening to what has been an incredibly helpful session and thank you to all our witnesses. Two questions still remain open for me, and to draw a conclusion I would like to ask both.

The first one is: in times of crisis, is the current process speedy enough to allow for information to be made available, before somebody gets a set of thumbscrews out? Is it possible for information to come out before the thumbscrews do? That is question one for me. Mr Horton, thank you, you have mentioned lessons learned from the Wakefield MMR situation—that independent scrutiny is really important and that concerns are raised.

My second question is: is research integrity permeable to a concerted, volume-based effort to obfuscate scientific research, almost like on the internet where you get distributed denial-of-service attacks where everybody tries to load the website so it falls over? Is the current research integrity system liable to that approach should somebody want to break the system's integrity?

Maybe I will start with that second one first and then I will come to Lord Ridley and Dr Chan. Can you be obfuscated by volume in these times of crisis? Mr Horton.

Richard Horton: Thank you, that is a great question. I can tell you from the volume point of view that it was huge. Everybody has rightly talked about, and it is much more important, the stress on the national health service in the past 18 months. If you look at the numbers of papers submitted to *Lancet* journals, they went up by factors of five or six. We certainly struggled to have resilient systems. It is a very fair concern.

However, I think that the broader media, especially social media, act as a very valuable corrective on what journals do and hold us accountable for our decisions in a very useful way. Let me give you a very good example of that.



Q171 **Katherine Fletcher:** The answer to that question requires social media to be available and free to use. Given that we are talking about Covid-19, I might ask you, especially as we are slightly over time, to stick to this. Was there a risk that Covid-19 research, especially in the early days, was obfuscated by volume, and does that special time perhaps make a special case, even if we have to redact it, at least to share the volume that you were receiving at that time and the source of the volume?

Richard Horton: Having lived through it, I can tell you we were certainly challenged by the volume but I do not think we were obfuscated by it. Our teams worked incredibly hard to try to sift out what was important and what was less important to publish. It was difficult, but I think for the most part we did a pretty good job.

Q172 **Katherine Fletcher:** That is fair enough. I ask the same question of Dr Chan or Lord Ridley.

Dr Chan: We have seen that happen. When groups are incentivised to promote a particular narrative, they can swamp the publication system. One very clear example of this was the pangolin papers. All of a sudden in February there were four pangolin papers put online at once and published in prestigious journals. One was even solicited by a prestigious journal. That fuelled media and public speculation that this virus came from a natural origin, from a pangolin. Meanwhile there was a paper that showed there were no pangolins or bats sold in Wuhan markets between 2017 and the end of 2019. That was kept under peer review and rejected by a journal and not shared. There is a lot of power in those journals to withhold and obscure information, whether intentionally or not.

Q173 **Katherine Fletcher:** Dr Chan, is that the basis of your point, which is that you need to know the full list, not just what was published?

Dr Chan: Yes, and that is why journals should mandate pre-printing before submission. Everything needs to be put openly before submission, not the day they are accepted after their revision. We want to see the original manuscripts. We want to see who wrote it and whether the data is available.

Q174 **Katherine Fletcher:** That was a very interesting point you made in quite a long list. I hope you will send that list to the Committee because my note-keeping was not up to speed.

Dr Chan: Yes, I will.

Q175 **Katherine Fletcher:** Lord Ridley, can we be obfuscated in our scientific endeavours?

Viscount Ridley: May I address your first question about speed, if that is all right? The pandemic showed that we could accelerate the scientific publishing process so as to get results out quickly. The Jun Zheng et al paper, which was pre-printed in January but with publication on 3 February, I think it was, was the first sequence of the SARS-CoV-2 genome and a comparison with the nearest bat sequence. It was very



important to get it out, etc., etc. It had problems. It left out crucial bits of information, as we later found out, but the point was that correcting the bad information ought to have been as quick as the original publications, and that is where the system fell down, so we got lots of stuff rushed into print that was then found to be inadequate.

The pangolin papers is another good example. Dr Chan wrote some very trenchant critiques of them—the inadequacy of their data, the duplication of their work and so on—and that got stuck in peer review for many months. We need a conversation to happen within the scientific literature much faster. Yes, somebody rushes something into print, but, yes, a rebuttal appears relatively quickly after that as well.

Katherine Fletcher: Thank you all for your time. I do mean it.

Q176 **Chair:** Finally from me to Richard Horton, talking about the speed and the volume of papers being produced, it has not escaped anyone's attention that Omicron is a subject of intense national interest. Just describe to me what research papers you are receiving and how you are going about making decisions to print and accept papers there. Given that we know the speed of transition of this virus is thought to be very rapid, it underlines especially some of the points that my colleagues have been making about speed. Tell us as the editor-in-chief of *The Lancet* what you are doing about this.

Richard Horton: First, in recent days there have been some quite frightening numbers that have come out of modelling groups in the United Kingdom. I understand that those numbers are the best that those groups can produce at the moment, but I also understand that there are critically important missing bits of information, and without that information it is very, very difficult to have reliable forecasts for likely numbers of infections, hospitalisations and deaths, because we just do not have reliable information on severity.

We are in touch with groups in South Africa that probably have some of the most reliable information at the moment about that. We have seen some early data on severity. That information is being written up and I hope that it will be submitted to us soon, but we have not received it at the moment. I hope and expect that we will be getting it within the next one to two weeks. Let us say that that information is successfully reviewed and published. Perhaps in the next three to four weeks they might present that as a pre-print—referring to Dr Chan's point.

We will have this information, I think, before the holiday, and then we can plug that into the models so that before the end of the year we will have a much better estimate of where we stand with Omicron. Right now we really do not have that rigorous and reliable information. On the basis of the numbers that I have seen from South Africa so far, they are still relatively small numbers, so again they are going to come with an uncertainty interval, which should give us a little bit of a pause before we are confident. I know this policy of planning for the worst and hoping for



HOUSE OF COMMONS

the best seems extreme, but based on the data that I have seen so far that seems to be quite a wise policy at the moment until we get better information.

Q177 **Chair:** Reflecting finally on what was said about publishing the whole volume of submissions that you have had for publication, given that even accepting papers as pre-prints involves a degree of scrutiny and analysis, which takes time, will you consider publishing as a web platform, if not in physical form, what you receive over the days ahead so that—you referred to social media earlier—armchair scientists can at least see the data, even if they do not have particular expertise?

Richard Horton: I understand. For about a year or so now we have had a pre-print server. When an author submits a paper to any of *The Lancet* journals we give them an option—we do not mandate it; this is Dr Chan's point that we should mandate it—that they can have it uploaded on a pre-print server so that it is completely visible to everybody. That promotes transparency and accountability, and people can see the original document as written by the authors.

That is going some way on what Dr Chan and you are alluding to. You may then come back to me and challenge me, "Well, why don't you mandate it?" The reason why is that we think authors should have the right to decide whether their work is in the public domain or not pre-peer reviewed. Sometimes in the health field they might have extremely sensitive and controversial findings and they simply do not want to make that work available until it has gone through a validation process. The authors have the right to decide that rather than we as editors having that right. But I accept that this is an issue for discussion.

Chair: You make a very clear case as to why there might be a reason why people want to have some sort of review before they are associated with it permanently.

It has been a fascinating session, and a fascinating conversation, on both the general aspects and the particular aspects of the origins of Covid. We are very grateful to Richard Horton, Lord Ridley and Dr Chan for appearing before us today. This concludes this session of the Committee, although our inquiry into research integrity and reproducibility continues.