

Science and Technology Committee

Oral evidence: [Research integrity](#), HC 350

Tuesday 21 November 2017

Ordered by the House of Commons to be published on 21 November 2017.

[Watch the meeting](#)

Members present: Norman Lamb (Chair); Vicky Ford; Bill Grant; Darren Jones; Stephen Metcalfe; Stephanie Peacock; Martin Whitfield.

Questions 115 - 276

Witnesses

I: Professor David Hand, Royal Statistical Society; Dr Damian Pattinson, Vice President of Publishing Innovation, Research Square; and Wendy Appleby, Registrar and Head of Student & Registry Services, University College London.

II: Dr Trish Groves, Director of Academic Outreach, *British Medical Journal*; Dr Elizabeth Moylan, Senior Editor for Peer Review Strategy and Innovation, BioMedCentral (representing the Committee on Publication Ethics); Catriona Fennell, Director of Publishing Services, Elsevier (representing The Publishers Association); and Dr Alyson Fox, Director of Grants Management, Wellcome Trust.

Written evidence from witnesses:

- [Royal Statistical Society](#)
- [University College London](#)
- [Committee on Publication Ethics](#)
- [The Publishers Association](#)
- [BMJ](#)

Examination of witnesses

Witnesses: Professor Hand, Dr Pattinson and Wendy Appleby.

Q115 **Chair:** Welcome, all three of you. To start with, would you briefly introduce yourselves and say where you are from?

Wendy Appleby: I am Wendy Appleby, and I am the registrar and head of student and registry services at University College London.

Professor Hand: I am David Hand, emeritus professor of mathematics at Imperial College. I am a statistician, and I appear here representing the Royal Statistical Society.

Dr Pattinson: I am Damian Pattinson, head of publishing innovation at a company called Research Square, which provides editorial support for publishers.

Q116 **Chair:** Thank you. There may be questions that you do not feel all of you have to answer. Do not feel obliged to answer every question if you feel you have nothing further to add.

Witnesses at a previous session suggested that the increasing rate of journal retractions is the result of better detection of research integrity problems. How can statistical methods be used to identify problems of integrity, including errors in calculations and dodgy datasets, and has the ability to do this improved in recent years? Can the process now be fully automated?

Professor Hand: The ability to detect problems with data in papers has certainly improved, which could explain the increased rate of detection and retraction. That could be a partial explanation; there may be other reasons as well. This cannot be fully automated. Detection of problems involves a subtle interplay between the objectives and the nature of the data and the kind of issues you are looking for. Given a particular class of problems, you could automate the detection of it. The software StatCheck does it for a very narrowly delineated class of problems, but in general the range of possible problems is so vast that it cannot be automated.

Dr Pattinson: I completely agree with that. You can do a lot of automation to assess whether, for example, there is enough information presented in a paper to show statistical analysis, but whether the right statistical test has been performed usually requires a level of expertise beyond what computers can do. It is the same with detection of things like plagiarism and figure manipulation; you can have automation to help identify issues, but it normally takes a person to decide whether it is genuinely a matter of concern or an innocent error.

Wendy Appleby: I do not have so much to offer on the detection side, but in terms of handling it I would concur with the other comments. We have had allegations brought to us at UCL. I am the named person in our research misconduct procedure, so I am here to talk about what it feels



like to be managing these things within a university. As Damian says, it is about individuals and expert judgment, and referring it to those individuals to make an assessment as to whether there was deliberate intent to deceive, or whether it was an error or misinformed.

Q117 **Chair:** As a quick aside, do you believe that every university should have a named person performing the role that you perform?

Wendy Appleby: Absolutely. One of the things we would suggest is that it be made clearer who those individuals are so that we can work with each other if we need to refer cases to other universities.

Q118 **Chair:** It could be published on the website and articulated.

Wendy Appleby: Perhaps it could be a list of the named individuals for the research misconduct procedure, hosted by UUK, UKRI or one of the other bodies, to make it easier for liaison between institutions; likewise with integrity officers.

Q119 **Chair:** Do you publish an annual list of the inquiries you have conducted as proposed by the concordat?

Wendy Appleby: Yes. We publish a report on research integrity. It is very important to understand that, particularly in the UK, there is a difference between research integrity as informed by the concordat, which is very much about encouraging and promulgating good practice in research, and research misconduct, which is much more about dealing with the problems. We publish a report on research integrity that covers all the good practice and the proactive things we are doing at UCL. That report also records the number of allegations we have had as concerns, and what has happened to them. We provide an annual report to our audit committee specifically on the research misconduct cases we receive.

Q120 **Chair:** Do you think that every university should be doing, essentially, the same as that?

Wendy Appleby: Yes, because every university is required to be signed up to the concordat. It is a condition in the HEFCE memorandum of association and agreement, and within the concordat is the requirement to publish a statement every year on research integrity, but understanding that research integrity is much broader than just misconduct.

Q121 **Chair:** What are the most common statistical mistakes that we come across in published research?

Professor Hand: Can I first make a general point? Most of the problems that lead to incorrect conclusions in scientific papers are due to errors or mistakes in the data, or the way the data are analysed.

Q122 **Chair:** Rather than deliberate misconduct.



HOUSE OF COMMONS

Professor Hand: Exactly. Obviously, fraud is particularly pernicious, but it is not the elephant; it is not the big cause. The most common causes would be oversights in pre-processing the data; ignoring missing values or inadequate ways of handling them; introducing errors when pre-processing the data, which happens quite often; and misunderstanding the statistical tools you are using. If you want it, I could produce a breakdown.

Q123 **Chair:** It would be very useful if you could do that. Incidentally, are you concerned about the level of errors that occur in published research that you come across?

Professor Hand: That is an interesting question. I strongly believe in the ultimate self-correcting nature of science. In the long run, errors will be resolved and made apparent, and the fact that papers are detected and retracted is an illustration of that; the so-called reproducibility crisis is a manifestation of it. In the long run I am not, but in the short run, and for the reputation of science and so on, clearly it matters.

Q124 **Chair:** It may relate to an insufficiently clear understanding of the use of statistics.

Professor Hand: Often it does.

Dr Pattinson: In my previous capacity, I was editorial director of a very large online journal, PLOS One. We saw a very wide variety of statistical errors, and those were just the ones we identified. The ones we saw were what some people call prosaic statistics. They are the very basic t-tests scientists perform to test their hypothesis. They are often very poorly performed or use the wrong parameters.

Q125 **Chair:** Does that alarm you?

Dr Pattinson: It does. It feels widespread, and I am not sure that peer reviewers always do a very good job at identifying errors. Personally, I feel there is generally a low level of statistical—

Q126 **Chair:** Incompetence.

Dr Pattinson: I would not go as far as incompetence; I would say the understanding is lower than it should be, across the board.

Professor Hand: I totally endorse that. More statistical training is needed at two levels. The first is for the individual researchers themselves. Although one recognises that they have to spend all their lives learning biochemistry or psychology and cannot spend too much time also learning statistics, nevertheless, that is one area where their level of expertise could be enhanced. Secondly, we need more professional statisticians who can join research teams.

Q127 **Chair:** As a collaboration.

Professor Hand: Exactly.



HOUSE OF COMMONS

Q128 **Chair:** From what both of you have been saying, it sounds as if there is an issue both with researchers and with peer reviewers. The errors or failures to understand can occur with both.

Dr Pattinson: We can perhaps go further into peer review later. For the majority of journals, peer review is essentially sending a manuscript out to someone with expertise in the area—you are not quite sure which area, or it may be only one part of the cover of the manuscript—and asking for their opinion. There is very little in terms of structure. Publishers are very nervous of asking too much of reviewers.

Q129 **Chair:** Is the process of peer review fit for purpose, and are there sufficient financial incentives for peer reviewers to do it properly? In other words, if it is something else they are trying to do in the course of a very busy life with their own research work, is sufficient time and attention given to the process, or do we need to rethink how the process operates?

Dr Pattinson: I do not know about time and attention. Certainly, reviewers are very busy people and they take it on as additional work that is unpaid.

Q130 **Chair:** Unpaid.

Dr Pattinson: In the vast majority of cases, it is unpaid. Essentially, it is work they do in the evening, for example. We know from surveying peer reviewers that they take it very seriously and spend a long time on it—many hours. They take it seriously. They do not always know everything about a manuscript, and because they are not always asked which areas they are competent in, it is quite easy for things to be missed out. Statistics is probably one of the worst examples. Some journals try to deal with that by asking reviewers whether they feel the statistics are appropriate. If the reviewers say they are not able to comment on that, you might go to a statistical review, but in the vast majority of cases it is just a matter of asking a reviewer what they think, and if they do not mention the statistics you assume they think they are fine, but because it is not properly teased out there is a risk of oversight.

Professor Hand: A reviewer will normally have just the paper, not the raw data, and will not be able to reproduce the analysis that the author has claimed to have done. In some disciplines, such as pure maths, the reviewers will go through the details of the maths to check that it is right, but in other areas such as medicine, in general the reviewer will not be able to check the analysis because they will not have the raw data.

Q131 **Chair:** It is quite a limited exercise. Do you have anything to add, Wendy?

Wendy Appleby: I have a slightly different perspective, thinking about it from the point of view of the responsibility of the authors of the paper—the researchers themselves. Something we have been considering at UCL is taking forward more information and awareness-raising of the responsibilities of authors for the journal article they are putting their



HOUSE OF COMMONS

name to. One suggestion we are very keen on, which we know COPE is very supportive of, is encouraging contribution statements. If there is an element in peer review where there is a serious concern, it helps for that to be dealt with from an institutional perspective.

Q132 **Chair:** Is there any protocol about how reviews should be undertaken, the standards to be met and the things you should be testing?

Dr Pattinson: It is very variable. Some institutions provide training for early career researchers in how to perform peer review, but many do not.

Q133 **Chair:** It is quite haphazard.

Dr Pattinson: It is. My feeling is that an awful lot is expected of peer reviewers now.

Q134 **Chair:** We could be dealing with something of profound importance—life and death situations.

Dr Pattinson: At times, yes. This is something publishers do not pay for. If they can get it out to peer review quickly, it is dealt with by someone else. As a result, publishers have perhaps been less willing to break out the things that perhaps are not suitable for a reviewer to look at. Figure integrity is a good example. Perhaps that is something the journal should be managing; instead, because of the opportunity to use reviewers to do it, they tend to take that option. A lot is expected of reviewers now, and that could be an issue.

Wendy Appleby: It is worth recognising—David probably experiences this—that researchers take it very seriously. They take the contribution to science and the knowledge base in science very seriously. I do not think we should underplay that, noting the comments Damian made.

Q135 **Chair:** But we cannot just assume that they are all performing to the highest of motives.

Wendy Appleby: No. Certainly in an institution such as UCL we would expect those standards. From a research institution perspective, engagement of staff in peer review is very healthy; it can help to develop the research culture within the organisation.

Q136 **Chair:** What are the limitations of statistical methods and software in detecting problems? Can excluded data-points, poor study design or bogus inferences be detected reliably, or is that expecting too much?

Professor Hand: To begin with, it is very difficult to generalise. There are also questions of definition, and that is why you need statistical expertise to look at it. I have a set of a dozen values; they are mostly single-integer values, one of which is 25. That looks anomalous. Does that mean I should exclude it? Does it mean that perhaps there is an underlying aspect of the mechanism generating particularly large values? Detecting anomalous values and what you do with them is intimately entwined with your objectives and the questions you are trying to ask.



HOUSE OF COMMONS

The short answer is that there isn't a simple algorithm that will solve all these problems. Obviously, it will require statistical expertise, but also understanding of the questions, the aims, the data and how they were generated, the measurement processes and so on. All of those are necessary and they will differ from study to study.

Q137 **Chair:** Are there any other contributions?

Dr Pattinson: I agree.

Q138 **Vicky Ford:** I have some specific questions for Wendy because of the experience that UCL brings to this. You are the named contact. Can you tell us a bit more about what that role is in practice? What does the job involve? In your standard procedures, is the investigation by the university itself or by independent people?

Wendy Appleby: That is quite a lot. If I forget to answer any point, remind me. First, on my role as named person, in an institution such as UCL research integrity is incredibly important to our reputation. Therefore, having a robust research misconduct process and procedures is also incredibly important. Within our procedure, as named person, my responsibility is to oversee the operation of the procedure. I make judgments in the initial stages of the procedure and help to provide advice on its operation, but I do not make judgments on the latter stages. In a moment, I will say a little bit about how the procedure works.

I am also responsible for making sure there are communications with other bodies and so on, if that is necessary, in dealing with allegations. One thing we do at UCL that we consider to be good practice is that my line manager is the vice-provost for operations, and the people who work on research integrity sit under the vice-provost for research, so we have a separation. I deal with research misconduct under more of a governance heading, as part of my responsibility.

Q139 **Vicky Ford:** It is a separate reporting line.

Wendy Appleby: Yes. It also means for staff that we hope that in approaching the integrity officer they would feel confident that it is not necessarily about misconduct; they might be seeking advice.

In terms of our process, we have a procedure based on the UKRIO standard practice procedure. We developed it in 2014. We had a procedure before, but we revamped it all, in the light of the guidance, in 2014. It is a three-stage process, which is standard in the guideline procedure. The initial stage of the process is a preliminary assessment, which is the stage I take. Effectively, it asks, does the allegation of misconduct fit within the definition of research misconduct, or should it be dealt with under a different process—for example, financial problems or a staffing process?



HOUSE OF COMMONS

If I decide that it fits within the definition, the next step is for it to go to screening. At UCL, we establish a screening panel, which is effectively a peer review. It is quite interesting to talk about this on the back of the peer review discussion. Typically, our screening panel is three individuals drawn from within UCL, but we are very careful to check that there is no conflict of interest with the research or the researcher where there is concern. One of the developments we have made in our procedure since 2014 is to have a pool of screening panellists. We use UCL's judicial institute to provide training for them, to help them know what they are looking for and what they are doing.

Screening is very much about saying, is there meat on the bones of the allegation? Is there prima facie evidence of research misconduct? The important thing to emphasise is that it is about an intention to deceive, because things can go wrong. The screening panel look very carefully at an allegation and bring an expert view, on a peer review basis, to determine whether or not there is evidence of intention to deceive. They have other options. They can dismiss it; they can route the allegation to a different procedure; they can also say that something has gone wrong, but there is not an intention to deceive, so it needs to be dealt with by other mechanisms, which might be further training, for example.

If it goes to the third stage, which is the research misconduct investigation panel, we establish a fresh panel of a minimum of three people, which will include an external member—I recall one panel where the membership was entirely external—and that panel will conduct an in-depth investigation.

Q140 Chair: You make a judgment on the make-up of the panel depending on the nature of the case.

Wendy Appleby: In forming the panels, we look for experts. As you have heard, it requires that level of expertise to get underneath the skin of the allegation and understand what is going on, particularly when we get to the investigation, and that level of the process.

Q141 Vicky Ford: In our previous session with UUK, they seemed quite relaxed about internal investigations and said it was in the interest of the university to protect itself and its reputation, but it sounds to me as if in your third stage you recommend having a bit of independence, so that it is not all self-investigation. Is that fair?

Wendy Appleby: It is not all self-investigation; indeed, in our screening panel, our procedure allows us to use an external member if we wish. It might be that we are seeking a particular form of expertise; it might be a very complex case. If we wish to, we can do it at screening level as well.

To go back to the UUK point, it is absolutely correct that the reputation of the institution—a university in particular—rests on the integrity of the research, just as it does around the academic standard of the awards and



HOUSE OF COMMONS

degrees it issues. It is the research side of the academic standards discussion.

Q142 **Vicky Ford:** The special inquiry that you conducted in one case suggested that the investigation should always be under an independent chair. How do you feel about that?

Wendy Appleby: One of the things we need to bear in mind in looking at the previous evidence session and the discussion on research misconduct is that we talk about misconduct investigations in two ways, and that is not always helpful. It can also get confused in terms of expectations of when we might engage with funders and other bodies for example.

When we talk about research misconduct investigations, we are referring to the entire three-stage process I described to you, but the actual investigation is the third stage. If the special inquiry recommended an external chair for the screening panels in complex cases, we would absolutely agree with that. If in our process we move to the third investigation stage, that is the body with the power to say that research misconduct is proven. The screening panel and I do not do that.

Q143 **Vicky Ford:** I suspect it is a very tiny number of cases.

Wendy Appleby: It is. Typically, we get about 11 allegations a year. Most things are dealt with at screening. That is why some of the developments we have taken forward with our process have been focused on making the screening more effective, and also speeding it up, because we have been criticised for not being as fleet of foot on screening. It is about a balance between the judgment being fair to the researcher and the speed of the process.

Q144 **Vicky Ford:** You are world-leading organisations. How do you think your processes compare with the rest of the world? Should we be looking at that?

Wendy Appleby: Practices vary enormously in the rest of the world. We have had engagement with colleagues in the US, for example. Interestingly, within the process in the US, which is much more externally regulated, institutional hands are tied a lot more. There is not so much freedom to look into problems, understand things, deal with a lot of things through screening and learn from screening and push that into integrity activities.

In the UK, we are doing reasonably well, but I do not think we should ever be complacent about these things. We have to take it seriously; it is incredibly important to the UK's reputation for science, particularly the reputation of universities.

Q145 **Stephen Metcalfe:** I would like to look at the specific area of image manipulation. Can you describe your experiences of image manipulation, and on what grounds someone might wish to manipulate an image?



Dr Pattinson: Perhaps I should explain it a little. When a scientific paper is written, the authors present figures, which are generally the most important part of the paper. That is where you present your data. That usually takes the form of graphical plots of the data and some examples of the sorts of things they have observed. In biomedical science, which is my area, it is primarily things such as micrographs and photographs taken through a microscope—blots, which are essentially precipitations of proteins or RNA, visualised on a screen. They come out, essentially, as little blobs.

Chair: A technical term.

Dr Pattinson: A huge amount of biomedical science is looked at in photographs of small blobs on gels. Those are the images that are most common in the biomedical literature, and they are the bits that are generally manipulated. It is important to point out that they are examples of a much broader dataset. Usually the graph will show the full data, or a representation of it, and examples of the actual observation.

Authors tend to touch up images fairly frequently. It is very rarely deliberate misconduct, but the rate at which figures are tweaked a little bit to make them look nicer is remarkably high. A study we did last year showed that about 18% of manuscripts had images that had in some way been tweaked. Other figures have shown a similar rate—about 20%. That is the general rate of what we call image manipulation. The vast majority of it is just adjusting the contrast a bit to try to make something a bit clearer. For example, if you are looking at a field of cells and the pieces you are interested in are at the far sides of the picture, you might try to condense the middle, for example. It is not deliberately misleading, but it is trying to create an example that is not represented—

Q146 **Chair:** It may or may not be. Presumably, you cannot tell.

Dr Pattinson: That's right. If you have chopped out the middle, for example, you do not know whether that middle had something they wanted to hide, or whether it was just a blank space they wanted to cover up. Standard practice is that you would make it clear that this is what you have done and you have separated out images. Similarly, it is very common that in these blots you have a few lanes in the middle that are not of interest, so you just splice them out.

Q147 **Chair:** Should there always be a statement such as, "We've adjusted the contrast," or whatever?

Dr Pattinson: Yes. As long as you say what you have done in any kind of manipulation you have performed, it is reasonable. General practice now is that you either make it clear with a line or you state in your legend exactly what has happened. That is the kind of issue we see.

Other things are duplication of parts of an image. For example, you may use the same control photograph in different places and say it is a different thing. Most of the time that is because people make mistakes;



they think it is from a different experiment. People have a lot of files and it is quite easy for that to happen. But there are a number of cases where the author may have rotated a piece of a figure and stuck it back on. In those cases, it is hard to see how it could have happened by accident.

Q148 **Stephen Metcalfe:** There is an intention to deceive.

Dr Pattinson: I think so. Whether that is because of any malicious fabrication is unclear, but for whatever reason, they want to present something and they do not have the example they want, so they tweak it to show something. A study last year of published literature suggested that about 2.5% of the manuscripts looked at that had these issues were of a kind to suggest that it was deliberate.

Q149 **Stephen Metcalfe:** That is one in 40, so it sounds quite a high figure.

Dr Pattinson: I would say it is very high.

Q150 **Stephen Metcalfe:** That is where there is deliberate intention to mislead.

Dr Pattinson: Yes.

Q151 **Stephen Metcalfe:** This is a serious problem. It is not something that should be dismissed.

Dr Pattinson: I would be nervous about the conclusions of a paper where you had seen those sorts of levels of manipulation.

Q152 **Stephen Metcalfe:** You mentioned biological sciences. Are there other disciplines that rely heavily on images, and are there particular disciplines that perhaps are more prone to image manipulation than others?

Dr Pattinson: I can speak only to biomedical sciences, but within that large area, anything to do with cell biology, molecular biology or anything where you are looking at those blots, seems to be particularly prone to those kinds of issues. I do not know about other areas of research.

Q153 **Stephen Metcalfe:** You said that 2.5% of papers had manipulated images that were not tidying-up exercises. How are those detected, and who should detect them? Who should be responsible for checking perhaps every image, or banning the manipulation of images, full stop?

Dr Pattinson: Who is doing it currently? The answer is that it is journal specific. Some journals do it very carefully and have a good reputation for doing so; other journals do not. The big problem is that you do not know who is doing it and who is not.

Q154 **Chair:** Should there be a standard, an agreed concordat or whatever, that sets the standard of what journalists should be doing?

Dr Pattinson: I would like to see more standards in this area. It has been overlooked for a long time. The scale of the problem has come to light reasonably recently. I think everyone knew that it was going on, but the extent of it is a fairly recent finding. Some publishers are taking it



very seriously. It is hard to know who is responsible for it. As I said previously, there is a reliance on peer review to pick up things like this. Peer reviewers are certainly not qualified to do so. I would say that academic editors-in-chief are rarely qualified to do so, unless they are particularly engaged and have complete understanding of the issue.

These are very complicated things that require a good eye and knowing what you are looking for, and that is not something a reviewer would have. Obviously, I have competing interests, in that we are trying to work with publishers to provide these sorts of services, but I would like to see more publishers doing it pre-publication.

Q155 **Stephen Metcalfe:** When you say a good eye, is that a good eye for spotting the tell-tale signs within an image of blurring, masking, cloning and so on, or the science?

Dr Pattinson: It is the former. Some people are just very good at spotting a pattern and seeing it somewhere else in an image. Duplication is the most common problematic one. Some people are just able to identify the same piece of image. We are working on software—others probably are as well—that can help with identifying pixillated areas, pixels that can be transferred elsewhere. There is some automation, but it is very early stages.

Q156 **Stephen Metcalfe:** At the moment, you rely on someone's eyes.

Dr Pattinson: Essentially, yes; someone spending half an hour staring very closely at these images. It is very expensive. You are paying someone \$20 to \$30 a paper and there are 3 million papers a year.

Q157 **Stephen Metcalfe:** There was talk about using filters and an automated system that might be able to spot this, but presumably it ends up being an arms race. If you are setting out with the intention to deceive, there are ways around them, aren't there?

Dr Pattinson: That is right. Any tools you develop you would not want to make available to researchers; it is something you would want to keep your hands on as much as possible. We have seen improvements in manipulation. For example, we now see fewer examples of splicing of lanes in gel blots, but we suspect we are seeing more cloning of areas, which is extremely hard to pick up with the eye.

Q158 **Stephen Metcalfe:** In combatting this, you have tools—someone with a good eye—but what would be the solution to weeding it out? Is it to ban image manipulation, full stop?

Dr Pattinson: We could certainly have clearer standards. There is definitely a need for better education on what is acceptable. How would we ban it? Who is ultimately responsible for it? The feeling is that it is probably the journals, but it is very hard to make any kind of policy across a wide and diverse range of journals. It would be very hard to do in practice.



HOUSE OF COMMONS

Rates vary significantly from country to country. From my experience, and certainly from the paper published last year, the UK is one of the best in terms of standards for image manipulation. There are very low rates of it in a comparison of papers submitted, so that is encouraging.

Q159 **Stephen Metcalfe:** Where is the highest?

Dr Pattinson: It is in Asia. We see very high rates in some countries in Asia, and that is something all journalists are aware of. As to who is responsible, that is clearly a matter for a panel such as this. Do we want to spend a lot of time dealing with how to prevent problems in Chinese labs, for example? That becomes a problem.

Q160 **Chair:** Presumably, there could be an international protocol that you expected journals to sign up to.

Dr Pattinson: Yes. If there were a standard that journals could sign up to, people would be supportive of it.

Q161 **Chair:** Are there any moves to do that? It is quite surprising that this rather haphazard landscape persists without anyone seeking to grasp the nettle and set standards.

Dr Pattinson: It is very expensive to do. Some publishers I work with question whether it is cost-effective to spend that money. An outlay of between \$20 and \$30 is significant when it is not quite clear what the repercussion is. To go back to your point about banning it, at the moment the standards are not there and a journal may have to retract a paper that has clear problems with images, but that is about as bad as it gets for them. Journals often feel that that is not enough of a threat to them to require a million-dollar investment in fixing the problem.

Q162 **Martin Whitfield:** You talked about the need for an annotation with regard to imagery. If the expectation was that there was something about the imagery and it turned out subsequently that it had been manipulated, that in itself is evidence that the value of the research is questionable. How hard would it be to achieve an academic understanding that all imagery should come with an annotation of explanation? It would not help to identify them, but when they were identified, if it came with a very simple statement saying that it was or was not manipulated, the group could take a view of the value of that individual's input.

Dr Pattinson: Who do you mean?

Q163 **Martin Whitfield:** Sorry, it was very complicated. If the researcher of the paper is required as part of the paper to say what has and has not been manipulated one way or another, and every paper had an expectation that that would appear in it, if it turned out subsequently that that statement was incorrect, it would allow people to raise very legitimate questions about the value of the paper and the value of that research and the research team.



Chair: We would know where responsibility lies.

Martin Whitfield: We would know where responsibility lies, and that, presumably, is not that difficult to achieve, provided there is a professional expectation that it appears in all papers.

Dr Pattinson: I think that is a good idea. Getting an author to say what they have or have not done, and putting the responsibility on them to sign something that says, "This is free from manipulation" is a fairly simple step that could be taken. Some journals do that in other areas—ethical oversight, for example. That would be fairly easy to implement, but the publishing landscape is very diverse and getting adoption across it is always very challenging.

Q164 **Martin Whitfield:** I wonder whether that could become a baseline requirement for subsequent citation of a report or research.

Dr Pattinson: Again, there is little standard in terms of citation. The problem is that there are no clear standards in publishing practice. There are certainly organisations—many are present in this room—that work in these areas. There is some kind of collective movement, but publishing is the most competitive aspect of the whole process, so publishers are nervous about making big changes or significant requirements of their authors that would increase the workload.

Q165 **Martin Whitfield:** Do you think it is a significant requirement to say whether something has been manipulated?

Dr Pattinson: In general, one thing journals could do to improve things is, as David said earlier, to require data to be submitted alongside a manuscript. If you had all the raw images that went behind the figures so that people could go back and look very easily at where they came from, that would hugely reduce this, but it is a lot of work for an author. When I was at PLOS, we tried a very strict data policy that said, "You are allowed to submit to us only if you make your data publicly available at the point of publication." It was quite an anti-competitive thing to do. We felt it was the right thing to do, and at PLOS we were in a position where we were able to do it, but it had some effect on submission rates. People were less keen to send to a journal where there were those requirements; certainly that was the feedback we received.

Q166 **Chair:** Have the requirements been dropped now?

Dr Pattinson: No, they are still there. It has done incredible amounts to increase the volume of data made available.

Q167 **Chair:** Do you stand by that?

Dr Pattinson: Yes, I think it is the right thing to do, but I can see why other publishers would not want to do it.

Q168 **Darren Jones:** I want to ask about training in the context of both researchers and assessors of research integrity issues. We have heard



that perhaps there is a lack of statistical training for researchers. I am interested in the level at which that training should be provided. I was an undergraduate human bioscientist, and I think I had one module on stats in three years. I became a lawyer so it has not hindered my career at all, but if I had become a researcher, clearly I would have needed more than a first-year module in statistical method. Is that provided to doctoral students, or at postdoc level? If so, why do you think it is not currently sufficient?

Professor Hand: I cannot make statements about what is generally provided, but further courses in statistics would normally be available.

Q169 **Darren Jones:** Are those further courses mandatory or voluntary?

Professor Hand: I would imagine they are mandatory, but there is the question of how much statistical education you can provide. You cannot turn them into expert statisticians as well as experts in whatever discipline they are in already; there just is not the time. Clearly, elevating their understanding of statistics cannot hurt and will help, but you also need experts in statistics itself to collaborate with them and spot the kinds of things they would not be expected to be aware of.

Q170 **Darren Jones:** At UCL, do you provide that collaboration between experts, statisticians and researchers?

Wendy Appleby: I am afraid I do not know the detail for that answer; I would have to get back to you. For postgraduate research students, training is an inherent part of a doctoral programme. At this moment, I am not able to comment on the level of statistical training vis-à-vis individual subject programmes, but I can certainly provide some information.

Q171 **Darren Jones:** Moving to assessors, I am conscious that this feels quasi-judicial in the processes you undertake when assessing allegations of research integrity. You said that at the preliminary hearing stage at UCL you decide whether something fits within the definition of research misconduct. What is that definition, and who decides it?

Wendy Appleby: It is not a preliminary hearing but a preliminary assessment: what is the allegation, and is this the appropriate process? In particular, if it is a small matter, or something where nothing further can be done—for example, a student who may have been failed for plagiarism—there is no further sanction that UCL can impose, apart from saying, “Don’t come back.”

In that process and in our definition of research misconduct, we have a procedure that is approved through our research governance committee and the appropriate framework, and we report as I described. Our definition is set within that. The definition is based on the model definition provided by the UK Research Integrity Office. If you look at procedures in UK universities, you will find fairly common definitions. One of the interesting things about the international picture, which Vicky



HOUSE OF COMMONS

asked about, is that within the UK our definition often goes much broader than perhaps you might see in the US, where they are very much focused on fabrication, falsification and plagiarism. We would include not getting appropriate ethical approvals, not following other relevant policies, gift authoring and a number of other things, such as inciting others to research misconduct. We have quite a wide-ranging definition to deal with allegations under our process.

I would not say it is a quasi-judicial process. There is judgment, and we have to make decisions, and in a sense that is partly why it is an internal process in the university. It is about our responsibility for the integrity of the research produced in our name, and making sure we have appropriate methods for dealing with concerns around that and that we bring to bear the appropriate judgments.

Q172 Darren Jones: On the point about whether it should be judicial, it may be semantics, but you have all talked about the need to understand intent. That is quite hard to understand.

Wendy Appleby: It is.

Q173 Darren Jones: In a legal situation, the prosecution would test the evidence so that an independent panel could decide, normally on the balance of probabilities, whether the intent was in favour of or against, ultimately, the defence. Do you think the process is rigorous enough?

Wendy Appleby: We use balance of probabilities as our threshold within the process. We are talking about the reputation of people and institutions. The people within your institution you might ask to look at a case in, say, screening, which is the most common thing that happens, have an interest in the institutional reputation as well, because their research is associated with UCL. But you are absolutely right that the intention to deceive is the key thing that distinguishes misconduct from other forms of failing. That is why the suggestion about having clarity as to whether an image is being manipulated, tweaked or enhanced is very helpful; it could help an institution in dealing with an allegation of that nature to understand what the researcher said about it in the first place, rather than what they might be saying now.

Q174 Darren Jones: What role does the person making the allegation play in that process? Do they put their allegations in front of the researcher, or is it left to you as the institution to decide with a one-sided argument from the defence that it is or it is not?

Wendy Appleby: It varies from allegation to allegation. We have had allegations around image manipulation, and they have come from a pseudonymous external person called Clare Francis, whom Damian may have heard of. The communications come in quite a vexatious style. A lot of the communications are published on the internet.

As an institution, we have to balance our duties towards a staff member and their and our responsibilities around their confidentiality and privacy,



particularly in the initial stages. We balance those things. In that sort of case, typically we would not engage with the person making the allegation, the complainant. However, typically it is another researcher who has come across something, be that somebody internal or external to UCL. We would then give the complainant and the respondent the opportunity to put their side of the case.

Depending on the complexity of the allegation, we would deal with the screening process either by email, because an expert panel can look at the evidence and so forth and come to a sound judgment by that means, or they might meet and interview the people. Consideration by email is something we have introduced in our process, because getting three busy senior academic colleagues, who are in great demand around the world and internally, to come together at a given point, along with the respondent, can be quite complex and time-consuming in itself, hence the email approach and having a pool of panellists readily available to us.

Q175 Darren Jones: You mentioned the training that the judicial institute provides. What type of training does that cover? Does it cover the final panel as well as the screening panel members?

Wendy Appleby: The training is geared to our pool of screening panellists, because in a sense that is a standing pool. A screening panel is typically three individuals, and we try to draw two from the pool. The idea behind it is that they absolutely understand the process and the judgments and decisions they are able to make; they know what to look for and they understand some of the underpinning principles of natural justice, fairness and so forth that are really important in dealing with these things. When academics who are experts are looking at a misconduct concern or allegation, it is hard to stop them wanting to go further, so the screening is to make a decision at a certain level, and then the investigation is a separate process.

Q176 Darren Jones: Do you provide training for the final panel?

Wendy Appleby: Very few cases get to an investigation panel; most are dealt with at the screening level. Each panel is different. We bring in external people and look for very specific expertise, because it is an important decision. It could make or break a researcher's career if there is a case of proven academic misconduct. We tend to approach it more in terms of briefing the panel and guiding them through officer support.

Q177 Darren Jones: The answer is that there is no training on judicial matters.

Wendy Appleby: There is no formal training as such, but we determine it on a panel-by-panel basis. We might ask our judicial institute to give them some briefing, or we might ask the person who is guiding the panel, or the chair if they are an internal person, to give the panel a briefing. I tend to step back from it a little bit, because within our process I must not interfere with the work of the screening and misconduct panels.



HOUSE OF COMMONS

Q178 **Darren Jones:** You talked just now about people having an interest in the reputation of the institution. If at the preliminary—I keep wanting to say hearing—

Wendy Appleby: Assessment.

Q179 **Darren Jones:** Thank you. If at the preliminary assessment and at the screening panel they are all internal people, is it not right to say that that leads to conscious conflicts of interest for those making assessments?

Wendy Appleby: You could make that assertion, but I do not think that is how it happens in reality. My role is to operate the process on behalf of UCL and to put UCL's interests to the fore. UCL's interests are absolutely in the integrity of its research and its good standing as a leading research-intensive university. The panels I have described bring interests around the scientific base, and actually screening panels are often horrified. I have noticed that in some universities' procedures they still refer screening to the head of the department where the allegation sits, or a single person. There is greater danger for conflict of interest there, but panellists tend to operate in very independent ways, and use the process and their expertise in forming judgments. In an internal employment process, say a disciplinary, a grievance or something like that, typically an internal set of staff would be involved in hearing that.

Q180 **Darren Jones:** The point I am going to in my line of questioning is that, because of the potential impact of research on the wider world outside the institution, surely this cannot be dealt with just like a normal internal disciplinary issue. That is why I have concerns. For you—not just UCL, but generally—there is an issue around the conflict of interests of those making decisions, the robustness of the process of assessment of the evidence in order to come to a conclusion of intent, and the way in which that stands up in respect of any decisions made. Do you and other members of the panel agree that these are more than just internal disciplinary issues?

Wendy Appleby: They absolutely are. When there is a decision that research misconduct is proven, it is made by an in-depth investigation panel that includes external members. That stage of the process absolutely includes external members, and brings that externality to it.

Q181 **Darren Jones:** Are there any other thoughts on that from the other panel members?

Wendy Appleby: When I was talking through this issue with a colleague at UCL before coming here today, he challenged me in a very similar way. He said we would not award a degree, be that a doctoral or an undergraduate degree, without an external examiner in the room. An external examiner nowadays is not judging academic standards, certainly in a taught degree, but they are looking at the overall process and its application. In a doctoral award, they are an expert and they make all the decisions around expertise. We talked about whether one of the options for the group UUK is setting up would be for the institution to have



HOUSE OF COMMONS

almost an external adviser who could bring in that expertise, looking much more at the operation of the process. I would hope that, if they looked at what we did at UCL, they would find it robust. Indeed, when we had our special inquiry, which was entirely external, they concluded that it was robust.

Q182 **Darren Jones:** Is the use of external members something UCL chooses to do, or is it commonplace across universities?

Wendy Appleby: For investigation panels, that should be commonplace, because it is within the UKRIO model procedure.

Q183 **Chair:** Do you think all universities should follow that approach, of including an external member on the panel?

Wendy Appleby: Assuming that all universities do, they should, but I cannot speak for others.

Professor Hand: I would like to make a comment in the context of annotating image modifications. Can I do that?

Chair: Of course.

Professor Hand: The same issue occurs much more generally. In an ideal world, an author would describe the modifications they had applied to the data when cleaning it. Data appear and you have to clean it to make it suitable for analysis and fitting models. You have to sort out missing values, check whether there are outliers and all sorts of other things, to make it suitable for analysis. In an ideal world, descriptions of all those steps and processes would be included in the paper, but in general they cannot be, because strict word limits are, necessarily and understandably, imposed by publishers on what appears in a journal. All that stuff vanishes, so it is precisely analogous.

Q184 **Chair:** But when the paper is submitted to the journal, that explanatory note could come with it.

Professor Hand: It could—as supplementary material or material subsequently placed on the web. That would be an excellent idea.

Chair: Surely, it should be a requirement.

Q185 **Stephanie Peacock:** I have a question for Wendy. You referred to this previously. Building on some of the questions Vicky asked, UCL was recently connected to a high-profile case of research misconduct relating to the work of Paulo Macchiarini. I know you touched on some of the issues, but could you briefly outline UCL's connection to the case and its response?

Wendy Appleby: Paulo Macchiarini was a collaborator of some of our researchers towards the end of the last decade—the noughties; indeed, he was appointed a visiting professor at UCL. His professorship was



terminated in 2014. Of course, what happened at the Karolinska Institute is well known and well rehearsed.

I have been at UCL for three and a half years. In my time at UCL, a number of allegations came in on the area of regenerative medicine research, focusing particularly on some of the sorts of methods Macchiarini was using, with slightly different angles in each allegation and lots of questions. One interesting thing about the way a research misconduct process works, or the stages within the overall procedure, is that it relies on allegation, a respondent and so forth. At UCL, we felt we had had a connection with Macchiarini, even though it was not current. We were doing research and working in the area and we had received a number of slightly different variants of misconduct allegations around the area.

We felt that we needed to step back from the approach where you need an allegation to look into a specific thing, and take a more generic approach, which was why we did the special inquiry, and that that should be independent. We had an entirely external panel for the special inquiry, with separate legal advisers we appointed and paid for, to help them in their work.

Within the inquiry report, there is a whole chapter on our relationship with Macchiarini. Fortunately, a lot of his latter work was not associated with UCL at all. It would appear from the findings of the inquiry that his main relationship pretty much ended in 2012. The inquiry team did, however, find quite late in the inquiry process, when rechecking all the transcripts and the documents around it, a very serious concern around the production of a device that was used in a human being outside the appropriate regulatory environment.

As a consequence, if you look at the recommendations of the inquiry, you will see that we were advised to report that as quickly as possible to the MHRA, as we did. We set about doing the inquiry because we wanted to allow for all the comments, contributions and allegations to be brought forward, and also because we wanted to understand what our relationship was.

Q186 Stephanie Peacock: The UK Research Integrity Office is a source of guidance. Did it support you in this at all?

Wendy Appleby: Not in this particular approach, because we felt that, as we are a leading research institution, we needed to show leadership and accountability for what happens in our name. We needed independent scrutiny, because it was an incredibly complex landscape with all sorts of things going on.

Q187 Stephanie Peacock: What are the broader lessons that you take from the situation?

Wendy Appleby: They made a number of very helpful recommendations; indeed some of them were around the operation of



our overall research misconduct procedure, and you can see them. Some of them were around scientific practice, and obviously working appropriately within the regulatory environment. I should add that the scientist who was found to have fabricated the device I referenced had previously been dismissed by UCL for financial misconduct, but we continued to investigate the issues.

There are lessons in terms of how we work with research councils—I know the Wellcome Trust is coming in later—balancing the interests of the individual and their rights and the expectations of funding councils, and the contract we have with them. Finally, within UCL itself, we are looking at the body of activity. It is a very wide-ranging area of research activity, and includes about 1,000 individuals, a huge number of staff across a wide range of organisational units. We were looking at whether we had the governance right, so that if ever there were a rogue action in the future we would have more robust oversight of it.

Q188 **Chair:** Has the Macchiarini case led UCL to reconsider related research into airway transplants, such as the planned Inspire and RegenVox trials?

Wendy Appleby: They looked in detail at those particular studies, but I would probably need to give you a written answer to that question. The inquiry team looked at those and has reported on them, but I cannot recall the exact detail.

Q189 **Chair:** I would be grateful if you could come back to us on it.

Wendy Appleby: Certainly.

Q190 **Chair:** There has been a written submission of evidence from Patricia Murray and Raphael Lévy¹ of the University of Liverpool that addresses lessons to be learned from this case. It sets out a number of recommendations. I would be grateful if you could give us UCL's response to the recommendations. On the point about the Inspire and RegenVox trials, they say very specifically that the proposal for funding for those trials included misleading information, which, in effect, still stands, and substantial sums of public money have been allocated to those trials. They say very clearly that the trials should be suspended. I would be interested to know what UCL's view on that is, and whether you feel you should be participating in trials where there are serious question marks. Bear in mind that patients died after that procedure was used. I am not in a position to judge all the reasons, but these are life and death circumstances.

Wendy Appleby: Absolutely.

Q191 **Chair:** It is important that we understand exactly what UCL's position is on that.

Wendy Appleby: I am very happy to provide that.

¹ Written evidence received from Professor Patricia Murray and Raphael Lévy ([RES0022](#))



Q192 **Chair:** The report of the UCL special inquiry says that communication with the Wellcome Trust as the funder of some of the research in question was “slow and lacking content.” What lessons can be learned about how universities interact with funders when investigating problems with research integrity?

Wendy Appleby: We need clear named contacts within the funding organisations, and a very clear understanding about what funders will do with information we provide to them, so that we can have those conversations with them.

Q193 **Chair:** Does UCL accept the finding that you were slow and lacking content?

Wendy Appleby: We accept that we could have been better—

Q194 **Chair:** Should have been better.

Wendy Appleby: We should have been better in our communications with the Wellcome Trust, and that is something we need to look at. A topic of discussion between funders and universities is about when you disclose to a funder an allegation of research misconduct. We need to be very clear about the stage for that and about what a funder is going to do with it. Understandably, if an allegation is to be dealt with at one of the earlier stages, and is not going to go through to proven misconduct, researchers are naturally concerned about their funders being informed of that. They have certain rights in terms of confidentiality as well. Clearer protocols and mechanisms for dealing with these things generally would be very useful. From our point of view, we need to make sure that we are working more closely with the Wellcome Trust, and being clearer about what we can and cannot disclose to them.

Q195 **Chair:** The KUH report—the Swedish report—was highly critical about patients being used for marketing purposes, and mentioned that the first patient operated on in Sweden was used to promote two universities, UCL and the Karolinska Institute in Sweden. What is your view about that? In the recommendations, it says that the practice of using patients to undertake publicity work should be banned immediately.

Wendy Appleby: Using patients would not be seen to be appropriate. In these sorts of studies, you get a report on the study, which can obviously inform the scientific community. Some of our colleagues report on operations, techniques, developments and so forth, but using patients to promote an institution does not feel appropriate, and we would want to support the recommendations of the Karolinska.

Q196 **Chair:** Reading through this critique, it seems to me that this case is an extraordinary case study in very poor standards of practice in all the institutes involved, and lessons must be learned from it.

Wendy Appleby: It was for that reason that we instigated the special inquiry. We wanted to understand what the lessons were for UCL, and what our involvement was. If you look at what happened at the



HOUSE OF COMMONS

Karolinska, the rector overturned one of the initial investigations into the actions of Macchiarini. Senior management at the Karolinska were complicit, which was why a number of them moved on. At UCL, we try to be very clear that the integrity of the research is our responsibility, and however difficult it is we need to deal with that. That is why we had the inquiry.

Q197 **Chair:** They think you should go further than the inquiry. They say that you should commission investigations similar in scope to those commissioned by Karolinska University Hospital to identify any shortcomings in how patients were treated and any breaches of medical ethics, including whether it was appropriate to treat patients under compassionate use, which, this document suggests, was a misuse of the exemption to the normal practice.

Wendy Appleby: The special inquiry looked at the ethical procedures around the operations that happened within the UK, particularly the ones associated with UCL, with our partner hospitals, and has reported on some of that.

Q198 **Chair:** When you come back to us, you can comment on whether you accept the case for going further.

Wendy Appleby: We can certainly give you a bit more detail around that.

Q199 **Chair:** The Macchiarini case cast doubt on the effectiveness of using stem cells to support trachea transplants, but a 2010 news article, still available on the UCL website, advertises Macchiarini's work and attributes the "success" of a windpipe transplant to the use of stem cells, which is clearly not the case. Does UCL have any procedures to retract its news articles and web pages following problems with research integrity? Is it reasonable to leave potentially misleading material like that online?

Wendy Appleby: As soon as I get back to UCL, we will be addressing that.

Q200 **Chair:** It will be removed.

Wendy Appleby: Yes. Thank you for bringing it to my attention.

Q201 **Bill Grant:** Can I ask about the responsibility for detecting problems with research integrity? Where does that responsibility lie? Should it lie with publishers or employers, and can we rule out peer reviewers who, I understand from today, are unpaid? Where does responsibility lie? I might add that I have sympathy for the peer reviewers.

Dr Pattinson: Publishers have responsibility for the integrity of the literature. If papers are published, publishers have responsibility for the integrity of the content they publish. If there are suspicions of misconduct, it is very hard for publishers to investigate those themselves. For example, they have no access to lab books or raw data. In those cases, it has to be sent to the institution for a full investigation. On



integrity, it probably lies with the institution, but, where publication occurs, journals have a responsibility.

Q202 **Bill Grant:** When you say an institution, would that be the employers? Is it one and the same thing?

Dr Pattinson: Yes. It is a bit of a grey area. There is an argument, which perhaps I will not go into here, about the rise of soft money, in terms of external funding as opposed to institutions, and that has perhaps led to a lack of clarity over who is ultimately responsible. If a researcher has all their money coming from external sources, who is the employer? Is it the institution or the funder?

Q203 **Bill Grant:** Are you learning towards the publisher? Are you hanging the hat on the publisher?

Dr Pattinson: Publishers could do more to show the value they add. The recommendation I would make is that publishers could be much clearer about what assessment they have performed on manuscripts prior to publication, so that readers can see very clearly what has been looked at prior to publication.

Q204 **Bill Grant:** The Publishers Association describe iThenticate—I nearly called it Authenticate—a plagiarism checker, as “a cross-industry initiative.” Did industry lead the development of that? Can we be confident in an industry-led system that is designed to detect image manipulation and identify databases that are too good to be true? Can we trust that system if it is solely industry-led, or should other people have been involved?

Dr Pattinson: The company that developed iThenticate went to Crossref, not to The Publishers Association. Crossref is essentially responsible for indexing the literature and assigning digital identifiers to it.

Q205 **Bill Grant:** Is Crossref across the UK?

Dr Pattinson: No; it is an international publishing organisation, which provides infrastructure for scholarly communications. By going to them, the company had access to the full corpus of literature and was able to index against the entire literature. It is a private company, but, because of its collaboration with that publisher-led enterprise, a level of oversight came into effect, which I think helped.

Q206 **Bill Grant:** We can be confident that it is what it says on the tin.

Dr Pattinson: I have been using it for many years, and it seems to do a very good job. Of course, false negatives are pretty much impossible to pick up, with plagiarism in particular, so there may be things it misses.

Q207 **Bill Grant:** But it is a good step forward.

Dr Pattinson: It is a great step forward. As I said at the beginning, it requires human intervention. It does not tell you whether something is plagiarised; it just highlights overlapping text, so there is a difference.



HOUSE OF COMMONS

Q208 Bill Grant: To come back to costs and responsibility, who bears the cost of checking papers for image manipulation and fabricated data before they are approved for publication? Would it be the research funder or the publisher? How might that affect papers that are submitted and reviewed?

Dr Pattinson: I am not sure that I understood the question. Who is responsible for papers prior to publication?

Q209 Bill Grant: Who bears the cost of checking the papers for image manipulation?

Dr Pattinson: My personal opinion—I am sure Catriona will have a different one—is that the weight of responsibility for the published literature should be on the publishers, and therefore they should bear the cost of making sure that the assessment is performed adequately. They should also be very transparent about what they have checked and make very clear to readers what has been assessed. I would put that on the publisher.

Q210 Bill Grant: We looked at various levels of allegations of fabricated data, plagiarism, manipulation or the intent to deceive. The highest one I picked up is an investigation where the allegations are taken seriously. As we heard, you put together a panel of experts. For me, and perhaps the whole Committee, the question is, what are the consequences for the individual or individuals who either admit it at that level or it is proven? It sounds like a court, but I know it is not intentionally so. Do they just drift away into the clouds? Do they go somewhere else? Are there consequences for those individuals?

Wendy Appleby: A research misconduct procedure is not an employment procedure. Because it is about looking as much at the research as the individual, we might be looking at research that relates to an individual employed by another university. For example, I dealt with a case referred to us from another university, because the research was undertaken at UCL and was in our name. We would let that university know the outcome of our process.

More generally, in terms of consequences, there would be a clear evidence base for a staff disciplinary process, with all the consequences that might follow. There is a series of actions in terms of rectification around the research. Retractions and so forth are picked up by the wider community, so there are consequences in terms of reputation, but it depends on whether they are an employee of yours, somebody else's employee, or still working. We look into allegations where some of the cases involving image manipulation go back 10 or 20 years and are now coming to light, but we will still look into them because of the history and the record. Of course, people have moved on; they might even have died.

Professor Hand: I would like to make a very quick comment on a couple of your questions. Wendy has just told you about the consequences



HOUSE OF COMMONS

within the organisation, but there are several cases of people who were found to have perpetrated fraud then being employed at another institution. There is not a single overall body running these institutions; they are very heterogeneous, so there is a serious problem.

On your question about who bears the ultimate responsibility, in my view it is the employers and the scientific community; it is down to them to enhance the notions of scientific integrity and what to expect from researchers. There is where real responsibility lies.

Q211 Bill Grant: The legal and medical professions have practising certificates. If there is an error on their journey in that career or profession, the certificate is withdrawn and they can no longer conduct their professional business. That does not apply to this type of business.

Wendy Appleby: No, unless there is some form of professional regulation otherwise associated with an individual's role. For example, if somebody was a clinical academic, research misconduct might raise questions about their professional registration as a doctor, but, more generally, there is not the same sort of professional regulation around researchers' work, or a register of practitioners, as you find in some of the other professions you identified.

Q212 Bill Grant: Chair, I apologise; this is my final question. If there was such a certificate, would it reduce the risk of malpractice, so to speak? Would it disincentivise fabrication and intention to deceive? Would it be an incentive for the researcher not to indulge in practices we would not approve of?

Wendy Appleby: Possibly. I am speculating, but given that the number of cases of actual misconduct and deliberate intent is small, I hope that it is just about rogue individuals and dealing with them appropriately. It is much more about the integrity agenda and encouraging good research practice for the rest of the body, which obviously is quite a key topic for this Committee.

Chair: Thank you all very much indeed. It has been a fascinating and important session, and we appreciate your time.

Examination of witnesses

Witnesses: Dr Groves, Dr Moylan, Catriona Fennell and Dr Fox.

Q213 Chair: Welcome, all of you. Thank you very much for your time in coming here. Perhaps you could introduce yourselves very quickly, just saying who you are and where you are from.

Catriona Fennell: My name is Catriona Fennell. I am here today representing the Publishers Association, and I work for Elsevier, which is a member of that organisation.



HOUSE OF COMMONS

Dr Groves: I am Trish Groves, director of academic outreach at the *BMJ*. I am also editor-in-chief of *BMJ Open* and deputy editor for the *BMJ*.

Dr Moylan: I am Elizabeth Moylan, editor for peer review strategy and innovation at BMC, which is part of Springer Nature, but I am here today to represent COPE, the Committee on Publication Ethics.

Dr Fox: I am Alyson Fox, director of grants management at the Wellcome Trust. We are a major funder of science research.

Q214 **Chair:** We heard in previous sessions about the rise in the rate of journal retractions. Do you think the peer review system is still fit for purpose if it is not catching problems at an early stage?

Dr Groves: It depends on what you mean by the peer review system. The traditional approach to peer review is that it all happens before publication and is mostly behind closed doors. People do not know who is reviewing their paper. There are many alternatives to that, many of which we have tested and now use at the *BMJ* and the *BMJ Open*. We have open peer review, which means that when a paper is published all the signed comments by the peer reviewers, and their conflicts of interest, are next to the paper for anyone to see, along with the author's responses and the original submitted paper. This system is also used at some of the BMC—BioMed Central—series journals.

Importantly, we now have post-publication peer review. This is particularly crucial when looking at research integrity, because people have the ability not just to send a letter to a journal, which it sits on for months and publishes a year later in print, but to have immediate, often moderated, online commenting. That may happen at the journal; it may happen at the large free public archive called PubMed Central, or on other websites, such as PubPeer. PubPeer was the website that revealed the very big scandal over stem cell research in Asia. Peer review is not one thing; it is not monolithic. I would argue that many recent changes have opened it up and made sure that it is a continuing process.

Q215 **Chair:** That suggests there is a variation in standard, because there is quite a variety now. Some very good things are emerging and happening, but that is not consistent or universal, is it? Is that true?

Dr Groves: Certainly, most health and biomedical research ends up in PubMed Central. That has a commenting function anyone can use, if they are a scientist and author themselves. It is not journal dependent; it is worldwide.

Q216 **Chair:** Do not feel that all of you have to answer every question, but does anyone else want to contribute to what we have talked about so far?

Dr Moylan: A lot of surveys have shown that peer review really does improve manuscripts and helps things get better. A lot of people value peer review. Peer review was never really designed to detect fraud, so



that is part of the issue. It was to see whether the science is valid, original and novel, so, if there is fraud, peer review might not pick it up.

Q217 **Chair:** But it is also important to pick up other errors, weaknesses and failures.

Dr Moylan: Yes.

Catriona Fennell: Although peer review was not originally designed to detect fraud, part of it was also a very practical limitation. For example, in the days of print journals, and less access to journals, it was harder to spot plagiarism. If somebody had plagiarised your paper and the article was sitting on a shelf somewhere you never saw, you did not know. One of the reasons why there was a serious rise in retractions was that people had more access to the literature, so the problem was more likely to be spotted in the first place.

From the publisher's perspective, to go back to something Dr Pattinson said, publishers do consider themselves responsible for detecting these problems. The Crossref-iThenticate initiative is one example of publishers working together very collaboratively to make sure that these issues can be detected. I think all publishers would say that the next challenge we consider very concerning is the level of image manipulation. The challenge there at the moment is a more technical one, but we are seeing very promising developments at, for example, Harvard Medical School and Humboldt University. There is a tool called ImageJ; there are also private companies working on it. The potential is there for technology for text, data and image mining to allow this. In a large survey of researchers, 80% hoped that peer review would be able to detect the problems. I am a bit more hopeful that the technology will help us get there.

Q218 **Chair:** You have touched on what I was going to ask. What is the role of publishers and funders in relation to research integrity? You have given a useful contribution. Are there any other comments?

Dr Fox: From a funder's point of view, research integrity is one of those activities and conditions where we rely on the institutions we fund to undertake in good faith that the whole of the research enterprise is absolutely dependent on the research being carried out well, with integrity, reproducibly and with no wilful misconduct. We are dependent on researchers' institutions doing that. The whole system relies on faith and trust.

We fund many, many organisations. In the UK alone, we fund 90 higher education institutions. As a funder, we simply cannot police all the grants we fund, or the institutions, so we rely on institutions to do that for us. Equally, we can help by making our expectations very clear. We, and any other funder, spend money. We are an independent charity, but other charities have donations, and research councils are reliant on public money. We want that money to be spent properly and well, to make the



best discoveries that help human health. We have a responsibility to make the expectations very clear, and, on the whole, we do. Funders get together and develop concordats, which we have done, to make those expectations clear.

Q219 **Chair:** I am not sure whether you were present for the first session.

Dr Fox: Some of it.

Q220 **Chair:** You heard the discussion about the findings of the special inquiry and that UCL had been slow to inform the Wellcome Trust in respect of their concerns. Do you want to comment on that?

Dr Fox: I gave evidence at that inquiry, and that is exactly what I said.

Q221 **Chair:** Do you feel you have reached a satisfactory position with UCL on the issue? Indeed, are you comfortable with where other research institutes are in terms of their responsibility to you to report things on an expeditious basis?

Dr Fox: We are comfortable with UCL in particular. We have received the final investigation report. We hope that UCL has heard our concerns, not about the conduct of the investigation per se but its relationship with Wellcome. You are right. It has been written in the report that UCL was slow. We are comfortable that institutions, in particular in the UK, the universities, are now very clear about their responsibilities in reporting any allegations of misconduct. We have heard that they are not entirely happy about the stage at which we ask them to report. We are comfortable with the stage at which we would like them to report, which is early. I have to say that that is unlikely to move.

Q222 **Chair:** Catriona Fennell, the Publishers Association wrote that publishers had increasing ability to monitor the performance of editors and reviewers and can use this to expose fraudulent behaviour. Could you give more details on what that means in practice? In your view, is the honesty and integrity of editors or reviewers a problem?

Catriona Fennell: In practice, it would mean being very clear that publishers have a strong policy of editorial independence. It is obviously important that the editor's decision making about the science is independent of the publisher. Bearing that in mind, when we talk about monitoring we are looking at a meta-level—for example, whether the editorial process takes a reasonable time. If it is too slow, some of the community does not appreciate it; if it is too fast, it might be a sign that it is not robust enough. On average, are papers being reviewed by enough reviewers? That could vary per field. For example, in medicine there tend to be more reviewers per paper, and very often there is a dedicated statistical reviewer. The problem with statistics was mentioned earlier. We see good practice in medicine, where they tend to send the paper specifically to a statistician. We are looking at overall patterns.



HOUSE OF COMMONS

In terms of reviewers, some of the behaviours we have seen emerge in the last couple of years are, unfortunately, when occasionally they abuse their position. For example, they use their position as a reviewer to get the author to add citations to their own work or to a certain journal. There have been a few quite difficult cases.

Q223 **Chair:** Are you suggesting that these practices are emerging and there is an increasing problem of deliberate malpractice? Is that your view?

Catriona Fennell: It is very hard to say.

Q224 **Chair:** You are just getting better at identifying them.

Catriona Fennell: It is very hard to say. For example, about 10 years ago, it became normal for most journals to have online editorial systems and for all of the process to be more structured and available. That has made it easier to notice these things, so it is hard to say whether they were happening in the past. Perhaps they were and we did not know. I do not think we can say. What I can say is that we are actively looking at analysis of these patterns for anomalies and people whose behaviour might be suspicious. To be fair, we believe they are a small minority, but we are looking at ways to detect it during the process, or after it has happened and then correct it. The tools are improving, but until about 10 years ago, when there was a more robust record of the process in editorial systems, it would have been quite difficult to detect those kinds of behaviours.

Q225 **Chair:** What do publishers do to ensure that their reviewers are equipped to spot problems with research integrity? Obviously, they have to behave properly themselves, which is the point you are making, but what is done to help them do their job properly?

Catriona Fennell: In most journals, there are two stages of what I personally would call peer review—you can debate that. I would say the editor's role is also part of peer review, but certainly the first stage tends to be the editor doing a check on the submissions. Speaking for Elsevier, but it is common among many other publishers, at that moment we would want things such as plagiarism checking and, when we have it at scale, image checking to be done, because we do not want to waste reviewers' time with papers that are plagiarised or have image manipulation. It is an unfair burden to place on reviewers. At Elsevier, we very much resisted moving those checks to reviewers.

Q226 **Chair:** You have done the screening before it goes to them.

Catriona Fennell: Yes; it is in the hands of the editor. It means that the editor has the tools. We talked about iThenticate. It is a good tool, but human judgment is needed, and it still takes five or 10 minutes for a human to check the report. It is not like a magic bullet where all the work is done; the editor has put a lot of effort into it.

Q227 **Chair:** Does practice vary quite a lot from one journal to another?



Catriona Fennell: It is becoming increasingly normalised. I believe about 4 million articles a year are now run through CrossCheck, or something of that sort; iThenticate has a couple of different names. Mr Grant had a concern that this was a publishing industry tool. The same system is used by most UK universities and many US universities, where it is called Turnitin. It has a different name, but it is the same tool. It is not unique to publishers or journals.

Q228 **Chair:** Are there any other contributions, or are you happy with what has been said?

Dr Groves: It is very important to point out that the role of editors at the screening stage is also to reject a ton of papers. You do not want to send out everything for peer review if you have no intention of publishing the article because it is not in scope or it is obviously flawed, inasmuch as it is not very well done and so on. Big journals might send only 20% of the submissions out to external peer review, having rejected 80% within the first couple of days. I want to flag up that, if any editor spots something they think looks like misconduct of some kind, they should not just reject the paper; there is a really important ethical obligation on editors to say, "Oh, dear," because they are in a position—

Q229 **Chair:** And raise it with the institute.

Dr Groves: First, you raise it with the authors. You do not accuse them, not least because they might sue you. You say, "There's something here we don't quite understand." Very helpfully, COPE has advice and even template letters on how editors can communicate with authors when they are first beginning to probe. It is really important that, if something does not smell right, you hang on to it; you do not reject it, because the minute you have rejected it you have no purchase over it.

Q230 **Chair:** Do you think the practice of journals in that regard varies?

Dr Groves: I am afraid it varies a lot. It would depend on the resources of the journal. Some journals have a very small editorial team; others have lots of full-time editors and many editorial support staff, and they are able more easily to handle the fact that we are largely rejection machines. The *BMJ* publishes 4% of the research sent to us every year, and that is pretty standard across the big journals. Most of our customers are dissatisfied, and we reject nearly everything.

Catriona Fennell: I am not sure whether you would concur, or have had the same experience, but we have noticed, when finding these problems, that some are clearly unethical and there could not be an innocent explanation, but there are also grey areas around plagiarism.

Q231 **Chair:** That is why questions should be asked.

Catriona Fennell: It becomes an educational moment, and that has a positive aspect, especially with a junior author, or if the author is in an institute that is not as well resourced as most of the UK institutes. The



HOUSE OF COMMONS

editor can provide an educational role. I think that has been a very positive side-effect, let's say, of checking for misconduct.

Dr Groves: Journals have another educational function. If you have detailed advice for authors and reviewers, it can itself be a learning tool. Some journals provide education. We provide online training. We also have patient review at the *BMJ* as part of our peer review process. We have pretty extensive education for patient reviewers online, with lots of support for them.

The other thing is the wider role of journals in publishing articles about methodology, research ethics and good practice. Journals can play an advocacy role to help change the culture as well as improving education generally. It depends on the journal, but certainly some of the bigger journals, in their fields, very much fulfil that role.

Q232 **Darren Jones:** My questions have two themes: one is the data flow around research integrity concerns between funders, institutions and publishers; the other is about training for the ability to spot research integrity issues at the publisher stage.

With the first set of questions, I am keen to understand how you get information from institutions on research integrity problems. Dr Fox just said that institutions were very clear about their responsibilities, but in the previous hearing, it seemed that UCL said that they do not know what happens with the disclosure of information, and whether they can be confident about that. That suggests that they are not clear and there is a problem between the two. Perhaps you would respond to that particular point, and then others on the panel could let me know how they get information about research integrity and what they do with it when they get it.

Dr Fox: With respect to the UCL comment, institutions should be clear about their responsibility in reporting investigations of misconduct. Whether there is a concern about what funders may do with that information is slightly different. It would be concerning if their concern about what we might do with the information prevented them from reporting, but our grant conditions and those of research councils are very clear that institutions must report any investigations. We ask that they do that at the screening stage, but we would not ask that they tell us about the person involved at that stage, because, as you heard, the allegations could be vexatious or malicious, or they could be frank allegations of misconduct, but they need to sort that out. If it progressed to full investigation, we would like to know.

Q233 **Darren Jones:** You want to know of all cases at screening stage, and the details of those cases if they get to a full hearing.

Dr Fox: Exactly, and we have made that much clearer in the past year.

Q234 **Darren Jones:** What do you do with the information?



HOUSE OF COMMONS

Dr Fox: Usually, we do nothing until the outcome of an investigation. We very much believe innocent until proven guilty.

Darren Jones: Good. That's encouraging.

Dr Fox: The consequences of someone being found guilty of misconduct can be rather serious; No. 1 being probably the researcher's reputation. Everybody in science knows everybody in science in their particular field, and they are dependent on their good reputations. Typically, if someone has been found guilty of research misconduct, we, as a funder, would no longer receive any applications from them for funding for life, because we think it is serious. That is what we do. We do not do anything until there is a finding; it is confidential.

Q235 **Darren Jones:** From the publisher's perspective, do you get information? How do you get it, and what do you do with it?

Dr Moylan: If an issue arises and is brought to a journal's attention, perhaps on a published paper, in the first instance we go to the authors and ask for an explanation, and we loop in their institutions. That can be quite tricky sometimes, because some institutions can come down on people quite harshly, and some institutions might not respond. As was said earlier, the publisher does not have the tools to do that investigation and the published article is, effectively, on hold until the investigation is completed.

That is where it is tricky, because publishers have a responsibility for the integrity of the published literature. What do they do in the interim? Often, people put an expression of concern on a published article or an editor's note, because they are waiting for the outcome of an investigation that might determine whether the paper is corrected or retracted. Then we get into the sorts of issues that are going on at the moment. How do you handle that? Should they have other terms? The publisher is waiting for the institution to get back to them.

Dr Groves: I concur. The traffic is that way; it is not the other way. It is not the institution contacting us to say, "By the way, we've discovered that a person who published 10 papers in your journal over the last 10 years has now done something fraudulent. You might want to look at those." It would be great if that happened; it does not.

Q236 **Chair:** That is worrying in itself, isn't it?

Catriona Fennell: I look at about 160 retractions a year, so based on that sample we have that experience.

Dr Groves: It tends to be journals. If one journal retracts, it often contacts other journals. I have been there 28 years. I do not know of a case when an institution has contacted us about that. Do you know of an institution that is doing it?



HOUSE OF COMMONS

Catriona Fennell: Yes. I see several cases per year. Most notably, very recently, there was a case that was unusual, because ultimately the editor and institution did not agree on the decision. That is unusual. I could come back to you with exact numbers. Maybe it is field specific or regional. I am not sure.

Q237 **Darren Jones:** It is interesting that in the field of medicine a journal the size of the *BMJ*, in 28 years, has never had an instance when a university has contacted you about an outcome on research integrity.

Dr Groves: We are often the ones banging on the door of the institution.

Q238 **Darren Jones:** I am interested in how different institutions act differently on outcomes. Elizabeth, you said that some come back to you and some do not. This might be a technical question, so don't worry if you don't know the answer. In a previous hearing, we heard about the concordat that universities signed up to. The Russell Group said that all its members were signed up to it and the UUK did not really know. Do you know whether the ones that do not respond are those that have not signed up to the concordat?

Dr Moylan: Often, these are cases that COPE hears. When a situation has occurred at a journal and they are trying their best to move it forward, they may bring it as a case to COPE. The COPE forum meets four times a year. Whichever COPE members have tuned in on the day can give their advice and comments, but often there are situations when something has happened and they are trying to reach out to the institution and not getting a response. They are stuck. What do they do next? Often, the advice is about whether there is a higher level they can approach. I would have to go back to COPE for exact figures. I do not think it is a particular problem in the UK. COPE is a global organisation, and it may just be variance around the globe.

Q239 **Darren Jones:** That is reassuring. Dr Groves, you mentioned—I think, as a side comment—the fear of universities or researchers suing you. Is that common practice with publishers? Do different universities act in different ways?

Dr Groves: It is an increasing risk. It is odd—maybe it is not odd—that, over the last five or 10 years, it is even scientists disagreeing with each other. Somebody publishes something and another group strongly disagrees with it, not necessarily because they do not like the science but because something about the overall hypothesis or interpretation is completely opposite to what they think. They often say, "I demand that you retract the paper or we'll sue you." You think, "Well, fine, go on then."

Catriona Fennell: Or they threaten to sue you for retracting the paper.

Dr Groves: Yes, it is very odd. It doesn't tend to go anywhere, but it is strange behaviour.



Q240 **Chair:** It is threats rather than reality.

Dr Groves: Yes. If you are trying to understand something in a paper that does not look right and you say straightaway, "We think this is fraudulent or deliberate manipulation of some kind," you are very likely not to get the same response you would get if you said, "Is it us? We don't quite understand this. Could you help us understand what went wrong? Perhaps you could send us more information and some of the background data. We would like to run this past our statisticians again in more detail," and that sort of thing. That is much more likely to—

Q241 **Chair:** —elicit a helpful response.

Dr Groves: Particularly if it is not deliberate misconduct.

Q242 **Darren Jones:** Does the *BMJ* have in-house lawyers?

Dr Groves: Yes. All publishers do.

Q243 **Darren Jones:** Do smaller publications not? Would other publications feel the pressure of threats of legal action more than perhaps the bigger ones like the *BMJ*?

Dr Groves: I suspect they might, but journals published by large publishing organisations, largely in the rich world, even if they are small journals of a big publisher, will have legal support at publisher level.

Catriona Fennell: In terms of the potential legal threats, I work for Elsevier, which is a large organisation. Our lawyers are very dedicated to this topic and spend a lot of time on it. Editors become very nervous, understandably, because they are concerned that they might personally be sued. I do not know whether there is any precedent for that. At Elsevier, we indemnify them so it cannot really happen, but it makes editors very nervous. It is a tough job and it does not help to have people sending you legal letters for making a scientific decision. I have a huge amount of respect for editors for having the integrity to stick to their decisions in that kind of stressful situation.

Dr Groves: And thank you to this House for passing the Defamation Act 2013, which exempted peer-reviewed scholarly communications from libel. That was enormously important and made a huge difference, because we were seeing this so-called chilling effect.

Catriona Fennell: Absolutely.

Q244 **Darren Jones:** Good. My second theme is about spotting research integrity issues. You said that it takes about five to 10 minutes of human judgment to determine whether it smells right. What is it that you are looking for in five to 10 minutes that suggests there might be research integrity issues? What does it look like?

Dr Groves: It is interesting that in the previous session you heard that from a statistical point of view often it is anomalies that raise concerns, but when you are reading a paper it is when it is too good to be true. If



something is perfect, it is probably not true because life is not like that, and research certainly isn't. What you want is a paper that explains in great detail the methods, not just the results, and then discusses the limitations; what went wrong, as well as what went right. That is science.

Q245 **Darren Jones:** In the previous hearing, we heard that peer reviewers, indeed editors-in-chief, are not qualified to spot research integrity issues. How would you respond to that?

Dr Moylan: It depends on the research integrity issue. We have heard a lot about iThenticate and plagiarism checks. That is using a tool. I also have the fear that you raised; when you bring in technology you have an arms race effect. We have seen situations where people have changed words in manuscripts to avoid detection, and it has been the human peer reviewer who spotted the research integrity issue and said, "This is plagiarised."

Catriona Fennell: By the way, it is not unheard of for the reviewer to be sent a paper that plagiarises the reviewer's work, because of course you send it to someone who is an expert. In the old days before we had software, that was fairly common.

Dr Moylan: There is plagiarism of ideas as well, which software will not pick up. It is the peer reviewer, who knows the community, the science and the research, who can point that out.

Catriona Fennell: I fully agree. The tools can help, but I cannot foresee human expertise being replaced by technology in my lifetime anyway. The technology needs to be analysed by someone with the context.

Dr Fox: Peer reviewers are absolutely the best people to spot fraud, if there is any. You are most concerned with plagiarism. Very clean data, for example, is a significant thing, and certainly image manipulation.

Q246 **Darren Jones:** My last question is on standards. Do you think there is any role for new standards across publishers, and maybe across funders through to institutions and publishers, if we are not being too grandiose, that could set a level playing field about how we spot these issues?

Dr Moylan: I think it is a cultural issue; it comes back to that. We all have our role to play and it is not one thing. I do not think the journals can necessarily bear it all. It is definitely the researchers, their institutions and the funders who can act together. If there is a culture of research integrity at an institution, which is nurtured—this is the way you are trained; this is the way you do things; this is the way research is run; we like to see reproducible research; keep your data—and if it supports institutions that organise the data, the funders can request data and the journals are empowered to put in datasharing policies. That is the way we will have to work together. That is partly why COPE is increasingly reaching out to institutions and has this year taken them as members, to start talking to one another and collaborating.



Dr Groves: What we are talking about are best practice advisory guidelines. They do not have teeth. COPE has many thousands of journals around the world in every discipline that are members. The guidance is great. It has developed from experience and an enormous archive of real cases, but it is advisory. The things that have tended to make a difference are the uniform requirements for biomedical manuscripts, which are now called recommendations. They come from the International Committee of Medical Journal Editors. It is a small but very influential group, because it is the top journals of medicine, including the *BMJ*. Currently, it represents 14 or 15 journals around the world, but everybody else says in their instructions for authors, "You must follow these guidelines." People tend to do that.

The other things that have made a difference are laws. For clinical trials, the requirement to register trials in a public registry before any patients enter the trial was driven by laws in the USA, which have been adopted, largely voluntarily, by other jurisdictions. As you may know, research fraud in the US has been treated as a criminal offence in some jurisdictions. Arguably, if you have misappropriated public funding and/or misrepresented science in such a way that it is harmful to society, patients or others, I wonder why it is not illegal in other jurisdictions.

Q247 **Chair:** You think that could be considered in this country.

Dr Groves: That is really a question for this House.

Q248 **Chair:** I am asking for your opinion.

Dr Groves: If you want global rules, they are much more likely to be led by laws, but I am not suggesting that the entire scientific enterprise becomes ruled by laws, because, as you can imagine, that would not necessarily be a wonderful thing. There is only so much you can do with voluntary guidance and good practice guidelines, of which there are many—COPE, the World Association of Medical Editors and so on.

Catriona Fennell: I would not call some of these things voluntary. You could say that there are three main points of hopefully positive influence on authors: their institutes, their funding bodies and journals. Researchers need those three bodies in order to do their research and move their career on. For example, the ICMJE brought out fantastic guidelines on authors declaring their conflicts of interest. This has become mandatory. If you submit to *The Lancet*, you have to send a stack of paperwork to show that every author separately has declared any conflicts and ethics approvals. If you cannot get funding or get published without doing something, people will do it.

Dr Groves: Yes, but it is not worldwide.

Q249 **Chair:** I am conscious that time is tight, but I have a quick supplementary. Dr Fox, you said that, if there were cases of misconduct, you would ensure that they received no further funding from you for life. Presumably, that means that you maintain some sort of list of people you



HOUSE OF COMMONS

will not fund again. Is that list made public? Clearly, there is a wider interest in ensuring that people who have been guilty of misconduct do not secure public funds, or indeed other funds that are available for medical and other research.

Dr Fox: It is something we have decided not yet to do, but it is something we could consider. I believe the NIH does that.

Q250 **Chair:** Do you accept that it is of wider interest than your organisation?

Dr Fox: It could be of wider interest. I would accept that.

Q251 **Chair:** People move from one institute to another; they could very easily emerge somewhere else and secure funding.

Dr Fox: They could. In truth, it becomes very clear on their CV if there is a gap in funding for some years because they have not been able to secure it, but you are quite right; it is something that could be considered.

We should not get the impression that deliberate frank research fraud is widespread. As a funder, we have not encountered that. Two years ago, the AMRC, charities, research councils, and the Wellcome Trust got together and decided that instances of deliberate fraud were rather small. At any one time, Wellcome has 3,500 active grants. In the past three years, we have had fewer than 25 reports of allegations. Clearly, we do not know what we are not hearing. Let's double it. It is still of the order of 1%, so that is rather small. Our concern, and possibly that of publishers as well, is more about the broader context of research integrity.

Q252 **Chair:** The standards of research.

Dr Fox: It is about standards and research design. Research practice has changed beyond recognition from the days when I did my PhD. It is so difficult to keep up with that. We feel it is much more vital for young researchers to be properly trained and for older researchers to be continually retrained. That is where I would like the institutions to focus attention.

Dr Groves: Falsification is a much bigger problem. We know that from clinical trials. We know that people change their outcomes and report only the findings they like. That sort of practice is extremely common, and in medicine it is potentially extremely harmful to patients. People do not see it as misconduct and do not consider it wrong.

Q253 **Chair:** Reporting only the findings they like seems to me to be straining towards misconduct.

Dr Groves: Me too.

Q254 **Martin Whitfield:** In the previous evidence session, we talked about whether or not researchers or research teams should, in essence, sign a disclosure confirming what they had done with regard to images and



HOUSE OF COMMONS

data. As a panel, do you think that would be a useful step forward, because deliberate fraud is a very small number? The choice of evidence seems to be a bigger problem, but, if there was open disclosure as to what was and was not used, it would allow people to reflect on the character of the researchers or the research team afterwards. What is your view about an agreed letter of understanding on the basis of the research?

Dr Moylan: A lot of journals require authors, through the submission system, to say what they have done, and that they have acted with integrity. I cannot remember the exact wording different publishers use, but there is an implication that you have ticked that box and you are not manipulating your images.

Catriona Fennell: Dr Pattinson mentioned something that I would differ from a little bit. He felt there were no real standards for what is allowed. There was, I thought, a well-known editorial in the *Journal of Cell Biology* in 2004 by Rossner and Yamada. It is very well cited, because it very nicely explains what is and is not allowed; hundreds of Elsevier journals, and potentially thousands of journals, follow it. I think there is a standard, and it is also part of author education. For example, if we do workshops in universities around the world, it is very clear as to what you can do and what you cannot do. I think the standards are there.

The only problem with asking people to declare the level at which they have manipulated images is that you cannot expect dishonest people to be honest. If they did something dishonest, I would not expect them to be honest about declaring it, and the honest people are already honest.

Q255 **Martin Whitfield:** I think the purpose behind it is not so much that they would or would not be dishonest, but, if it subsequently turned out that there was dishonesty, it would be reflected more severely when someone had assured everyone that they had not done it, rather than standing there and saying, "Oh well, it was really just a mistake." There would be an assurance. I wonder whether the same applies for the data.

Dr Groves: For several years, every research paper in the *BMJ* has carried a statement by the authors that says, in effect, "We solemnly swear that everything in this paper is true and accurate and reflects the data, and we declare that any changes we made from what was planned have been explained in the paper." It is a transparency declaration. All authors have to agree to that. It is published with the paper as part of the manuscript. A couple of other journals have picked it up as well.

Q256 **Martin Whitfield:** The word limit on a journal does not prevent that.

Dr Groves: You do not need word limits. Word limits are a print thing; we do not have word limits for our researchers.

Catriona Fennell: Word limits are increasingly a thing of the past. It was very much a print thing.



Q257 **Martin Whitfield:** There has obviously been an increase in the number of journal articles that have been retracted for a variety of reasons—detection awareness, the temptation to cheat, or indeed, as you mentioned before, problems with the actual study design. Would you like to comment on the study design element and the problems that is showing at the moment?

Dr Moylan: There are many reasons to retract an article. Retractions are not necessarily bad. A lot of stigma is attached to that. If something has gone wrong with the experiment, or, oops, the code sampling was wrong and not quite what somebody anticipated, the publisher has a duty of care to make that correction or retraction as they see fit. It is not necessarily all bad. The way research operates is inherently messy; mistakes happen. We have to be comfortable with that. We are all human. How we fix it and make that transparent is the key.

Catriona Fennell: You mentioned in particular study design. There may be something worth mentioning from the perspective of publishers and journal editors—I should have said that I also speak for editors. One example is the initiative of Chris Chambers of Cardiff called Registered Reports. I understand that it does not apply to every field, but 80 journals globally have taken it up. It tries to separate the design from the results.

There are other initiatives. In psychology, the journal *BMC Psychology* has something called Results-free Review which is similar. Basically, you review the paper without the results, so you focus on the design. Increasing importance is being placed on methods; for example, Cell Press has STAR methods, which put way more attention on method. There is an effort to try to look more closely at the design, even in Registered Reports, before there are any results. It is very positive, because reviewers have the opportunity to influence the actual research before it happens. One of the jokes you hear from authors is that, after they have spent three years working on something, there is always some reviewer who will tell them, “Oh, that would be great, but could you go off and do another six months’ worth of experiments?”, which is not what they want to hear. If there is a problem with your design, it is a bit late to hear about it a few years later. If you are told the feedback earlier, that is a very positive development.

Cortex, which is published by Elsevier, was the first journal to launch that four years ago. The uptake is not huge. It is a high bar, and I do not think Dr Chambers would say anything else. The intention is for it to be a high bar. One of the wonderful things about science and scientists is that they set a very high bar for themselves. The scientific community does not say that research integrity is just the absence of misconduct. They are talking much more broadly; they are saying it also comes down to reproducibility. That is difficult; that is setting a high bar. I respect the community for setting that bar, and publishers try to help them to reach it, but let’s be clear that it is difficult. Let’s also be clear that, largely,



HOUSE OF COMMONS

science is moving forward. It is applied. It works, to a large extent, so, although there is clear evidence of problems around reproducibility, I am also hopeful. We still believe in science and we believe in the process, and it is great that it sets such a high standard for itself.

Q258 Martin Whitfield: I will mine into my question a little bit and see whether there is any disagreement as I go there. There are more historical papers being withdrawn as we move forward. Is this a bubble that we will come to the end of, or will it continue? The reason I ask this is that other papers have cited subsequently withdrawn papers. It is a huge discussion in its own right, but I am interested in who is responsible for chasing those references and saying, "Oops. If there are problems here, are there not problems with your paper?"

Chair: Keep your answers short if you can. We are tight on time.

Dr Groves: We know that this is a massive problem in medicine; hundreds of thousands of patients might have been affected by papers that continued to be cited after they were retracted, some of them for fraud. It is a massive problem, and it is one of education at the institution initially. We know anecdotally that a lot of people put references in their papers without actually reading the papers they cite in the reference list. They have not bothered to check at the journal website or in an index, such as PubMed or MEDLINE, that it has a big thing that says, "Look out. Retraction." They do not check. That is initially the responsibility of authors. Some journals have systems where, when a paper is to be published, during the technical/copy editing phase all the references are checked. At that point, a good copy editor ought to pick it up and say, "Hang on. This one's been corrected," and it should come back to the handling editor and the author, but I do not know how often that happens.

Dr Moylan: I wonder whether anyone has done a study that shows that people could be citing a retracted paper, but they might be citing it in the right way, in full knowledge that it has been retracted. They know that and that is why they are citing it, because now we are following a different avenue.

Catriona Fennell: A certain percentage definitely cite it in the sense that it is a controversial retracted paper, or something like that. There is a lag. If a paper is retracted, papers may have already been written that cite it; they are in the editorial process, and do not come out for maybe five or six months. It could be that the person was not aware of it at the time they wrote it, and we would hope to try to catch it in the editorial process. After about a year and a half, if I remember the data, you see the citations drop off, because it becomes well known that the paper is retracted.

Q259 Martin Whitfield: But there is a problem with historical papers because they have been out there much longer, with the veneer of authority, to be cited.



HOUSE OF COMMONS

Catriona Fennell: You mean they have already been cited many times before they were retracted.

Q260 **Martin Whitfield:** That is what I was thinking. Does anyone have responsibility to chase those up?

Catriona Fennell: It is a collective responsibility, and certainly for the authors. There is a lot of education on good scientific writing, and a big part of that is proper citation; it is fundamental to use solid foundations. It is also for publishers.

Q261 **Martin Whitfield:** Do authors have authority to take care of their reports for the entirety of their life?

Catriona Fennell: Retrospectively?

Dr Groves: One would hope so.

Catriona Fennell: It is something that pretty much all publishers have tried to be more and more transparent about. If a reader links to the retracted article, the reference is very visible; the authors who send us legal letters say it is too visible. It has "Retracted" across it in big bold red letters. I think all publishers are moving in the direction of making these things more visible.

Dr Groves: If the question is whether this will become more common, the answer is yes, unless we change the academic reward system. While the pressure is still to publish in journals with a high impact factor, which we know is the rate of citation, the signs are that it will continue. If you shift the system to rewarding things such as making your dataset available—"Look, great, 10 other people have done research using my dataset"—that deserves promotion.

Q262 **Chair:** You would strongly advocate reform.

Dr Groves: Absolutely.

Dr Fox: As an example, we ask only for your top 10 or 20 papers in an application, and even that we are now changing to say, "Tell us your most significant research outputs," to try to get away from the obsessive focus that a published paper in a high-impact journal is the measure of success for a scientist.

Dr Groves: It is a real impact.

Chair: You are all talking far too much.

Q263 **Stephen Metcalfe:** You have actually answered my question without me wasting time asking it. It was about pressure to publish and whether that is a motivation. I am very pleased to hear that that is changing. I will not cover that ground again, but I will ask about reproducibility. Some people have described a reproducibility crisis. Some put that down to the fact that it is not very attractive to redo work someone else has done; it is not



HOUSE OF COMMONS

going to get published and you do not get much credit for it. How much of a crisis is there, and what is the way round it?

Dr Groves: It varies from field to field. We have lots of evidence from psychology. I know Professor Dorothy Bishop gave evidence here recently. She has led the way on initiatives to stop it. We know there is a problem in quite a lot of basic science, and Registered Reports is one way of stopping that. As far as we know, it varies a lot from field to field.

Dr Moylan: Funders may have a role to play. That is the case in the Netherlands, where huge grants are given to people to show that they can reproduce work, so maybe that is also the way to go.

Dr Groves: It is very important in clinical trials. We have the whole Cochrane collaboration whose job it is to pool trials, put data together and discover the state of the art for a particular research question. Reproducibility is high science; it is not something funders should not consider worth funding and supporting. By the way, the system also needs to support the data scientists, technical platforms and so on that allow proper reproducibility to happen. A big investment is needed in that.

Catriona Fennell: It is important to make a distinction between reproducibility to some extent and research integrity; it would be damaging if there was a perception that, because your work could not be replicated, you did something unethical. That could be the case for a very small percentage, but normally it is not. It could be for another reason outside your control; it could be an antibody that was not as stable as you would like. Some of it could be very much improved with education and more focus on transparent methodology. We heard earlier about the importance of statistics. For example, in psychology, statistics are crucial, and there is a large focus on improvement. It is important to show that there is not necessarily a direct connection with misconduct.

Q264 **Stephen Metcalfe:** I accept that, but do you see the unwillingness to publish reproduced research as an issue or not? You are shaking your head.

Dr Groves: Not at all. It is not a problem any more. You may not get into the very top journals, because their job is to publish stuff that is brand-new, but, for the zillions of other journals, reproducibility studies are incredibly important and very widely published, as are studies with so-called negative results.

Q265 **Stephen Metcalfe:** You touched on open access publishing or open datasets. Do you think that would be a positive contribution? Is there any further comment you would like to add?

Dr Moylan: If people are transparent about their data, it can head off problems later. I was talking to somebody who went through an honest error retraction, Richard Mann. He has a blog on it. He said, "If only I had



HOUSE OF COMMONS

shared my code and my data, it would have solved the problem.” He is now a believer in that as well.

Dr Grove: It is open methods too, not just open access and open data. The methods are absolutely crucial.

Dr Moylan: Open research.

Dr Grove: Yes.

Dr Fox: It is increasing massively.

Catriona Fennell: It is somewhere that certain journals can also help to have a positive influence. Somebody mentioned word limits earlier. In the past, with print, if the author had a word limit, there was a risk that they would reduce the method section, and you basically ended up with somebody writing a recipe: “Throw some flour in a bowl. Add some butter and throw it in the oven for a while.” Actually, what you need is what type of flour and how many grams, what temperature should the oven be, what type of butter, and so on. This is definitely a positive direction, but we also have to reward behaviour.

There is discussion about the fact that trying to get into very influential journals could have a negative force, but there is also a positive force. Top journals can require people to do things that they may not do for, let us say, a more average journal, and, when top journals do it, they normalise it and then the other journals can follow. We see an example of that with datasharing. The *New England Journal of Medicine* has just made a strong statement about datasharing. It may not happen overnight, but I am sure that, in following years, we will see that other journals want to imitate the top journals, so it will become a stamp of quality to have high standards and methods, declarations of interest or whatever. It will become synonymous with the top journals. I think that is a positive aspect of those journals.

Dr Moylan: Especially when funders mandate it.

Q266 **Martin Whitfield:** To pursue that point slightly, we heard a lot of concern about the level of statistical knowledge among researchers. A number of witnesses have come back to problems that occurred because of fundamental misunderstanding of statistics. That is clearly important, but do you think that algorithms and IT will save us from that potential statistical nightmare, or is it something that has to go back to fundamental teaching, all the way back to high school, or maybe even earlier than that? Does there need to be a far better real-life understanding of statistics? Do you think that a statistical expert should be part of every research team almost by default, where their absence is the exception rather than the rule?

Dr Fox: I am not sure that is realistic. Statistical training for young researchers is probably quite poor, inconsistent at best. It is not just stats per se; stats cannot answer everything, but the basic fundamentals



of designing an experiment and research programme are really hard. Traditionally, but not exclusively, a young researcher coming into a lab to start a PhD, or their first postdoc, is taught by the person above them, so it is generally not even the PI—the lab head; it is the postdoc or senior postdoc. In bad cases, essentially it is the blind leading the blind. Stats is really hard and research design is increasingly hard as the complexity of research increases manifold. It goes right back to training at the beginning, but it cannot be just a one-off.

Often, institutions are getting much, much better at recognising this and putting in place training programmes for young researchers, but it is usually one-off; it has to be repeated every year. A lot of lessons could be learned from industry—pharma—where the rigour of record-keeping, for example, is so much greater than the rigour in academia, perhaps because the stakes are higher. Ultimately, their records will go to regulatory authorities that pore all over them, and that includes the stats tests used for each individual experiment.

Q267 Martin Whitfield: Do you think funders have a responsibility to demand a higher level of statistical design within the format before you fund? Do you have a role in that?

Dr Fox: I do not think we should demand such specifics about statistical design, but we can and should demand the highest rigour of research integrity, part of which is statistical design.

Dr Groves: It is relatively unusual for statisticians to be part of the grant-giving panels at funders. I would argue that they absolutely have a crucial role there. I know it happens sometimes at Wellcome, but not always.

I want to flag up one interim thing, which is between the institution and the journal. It is a great movement to produce templates for how to write papers. They are evidence based; they are based on people doing very detailed literature reviews, working out the best way to report a particular paper. The best known is the CONSORT statement for writing up a randomised control trial. It walks authors through everything that ought to be reported in the paper. It does not say, "You have to do it like this," but it says, "Did you do it? Tell us what you did." At each part of the paper, it prompts you to report the important methodological points and anything important about reporting the data. That is a really important move. It is burgeoning throughout medicine in different study designs, and it is beginning to move into basic science as well. You may think it is cookbook writing, but there is pretty good evidence that it is not just improving the quality of papers and their readability; it has an educational function. People say, "I didn't do it this time, but I'll make sure I do it next time."

Catriona Fennell: There are some nice tools, but I do not think tools will solve it on their own. There are some nice supporting tools, such as Penelope, a tool from James Harwood, here in the UK. It goes through



HOUSE OF COMMONS

the manuscript, text-mines and checks some of the statistics, which is great. It also checks whether it is compliant with the ethical standards. Have they made the declarations they should have made? For example, is there a link to data and those kinds of practices? Technology can help, but it is part of a bigger culture. It is not something on its own; it is not a panacea.

Q268 **Martin Whitfield:** There needs to be an ethos; it needs to be based in ethics, as I suppose research always has been. Do you have anything to add, Dr Moylan?

Dr Moylan: I think it has all been covered.

Q269 **Bill Grant:** It was mentioned earlier that COPE is minded to bring together institutions and journal editors to tackle the obviously small number of research integrity problems. Where do you see that going? What would you hope to achieve from it? Is it a short-term thing, or does it have longevity?

Dr Moylan: It is a pilot, which is an initiative in the right direction. Journal editors and publishers often bring cases to COPE, and the No. 1 category of cases that is constantly rising are authorship disputes. If there is an authorship dispute on a paper, there is not a lot that the publisher of the journal can do about it but refer it back to the authors to talk about and bring in the institution to say what went on. Perhaps we can prevent those problems occurring more upstream.

Q270 **Chair:** You mean plagiarism issues.

Dr Moylan: Not plagiarism; authorship issues.

Dr Groves: Infighting between authors—for example, “I refuse to publish this paper until this person’s name is taken off.” That sort of thing.

Dr Moylan: There may have been a gift author down the corridor on the paper who should not have been on it, or a ghost author we do not know about. Those are the sorts of things that come up. If we can, we should address those upstream of the journal, where they are taking place, and work on clear guidelines for authors, authorship contributions and what people did, and make that talked about. That case was just one example. We write guidelines for peer reviewers so that editors can share them, but peer reviewers are researchers, so we should talk to the researchers. The researcher is the peer reviewer and the editor all wearing different hats, so we absolutely need to be having those conversations.

Q271 **Bill Grant:** You are enthusiastic about the pilot and hopeful about it.

Dr Moylan: Yes.

Q272 **Bill Grant:** Excellent. How do journals and research funders work together more generally to explore the small number of occasions when there are questions about research integrity? Do funders work together—

Dr Fox: There is a bit of a gap.



Catriona Fennell: In certain areas, they have been collaborating more.

Q273 **Bill Grant:** I do not want to make you all collaborators—that sounds like a wicked word. Working together, maybe.

Dr Fox: We are at different ends of the research enterprise. We are right at the beginning, as they come in with their ideas, so we are increasingly less bothered about authorship because we want to focus less on publications. If there is a retraction from a journal, we do not share information very much. With the increase in technology, given that the grants are acknowledged at the end of every paper, it is something we would probably be able to fix.

Catriona Fennell: Publishers are looking to extract from the articles which funding body they came from, so it is becoming more and more possible to do that.

Dr Groves: We have occasionally had cases where we have gone back to the funder—it might be the institution, or a hospital or health authority; it might not be a grant-awarding charity or national funder—but crucially to the ethics committee that approved a terrible study, because that is a very important feedback loop. We have done it sometimes, but it is not routine.

Q274 **Bill Grant:** But there is room for improvement.

Dr Groves: Yes.

Dr Moylan: Yes.

Q275 **Bill Grant:** As a lay person, I suspect it is a long journey from Petri dish to publication, and that journey will have a number of players. Does anybody recognise any other country that brings these players together better than we are doing in the UK? Is there any exemplar from a global perspective that is doing better? I am not suggesting plagiarism.

Dr Moylan: I think the COPE written evidence to this inquiry mentions that Australia has a research funding body that represents the academic institutions in Australia, together with the universities of Australia. They place a heavy influence on research institutions to foster a culture of research integrity. They require reporting of cases of research integrity at funding level, so maybe there are some lessons to learn there.

Q276 **Bill Grant:** It is not the complete model, but there are lessons to be learned from it.

Dr Moylan: Yes.

Dr Groves: The US is pretty good. If it is research funded by the National Institutes of Health, the processes are pretty robust and thorough. That is the place you go for best practice; they have the best advice on image manipulation, for instance. They have a fantastic website full of tools, ideas and policies, e-learning, you name it. That is a great



HOUSE OF COMMONS

resource, but that is partly about protecting public money, so it is very important. They do a good job with NIH-funded research.

Catriona Fennell: I agree that the NIH standards are very clear. They have mandated that clinical trials must have a data management plan, and that is an area where they are very strong. I do not know what percentage of US research comes from there.

Dr Groves: The National Institute for Health here and our Health Research Authority are also leading the way on many other things. For instance, the HRA is the only body overseeing ethics committees that mandates that you have to register a clinical trial as a condition of ethics approval for that trial. It is the only one in the world at the moment. It is leading the way, so there is some very good practice here too.

Chair: Thank you. We have reached the end. We really appreciate your time. It has been a fascinating session. You have all spoken far too much and played havoc with our timings, but we appreciate your being here and sharing your thoughts with us. Thank you very much indeed.