

# A non-anonymised review of annual statements on research misconduct from UK institutions in 2023-4, with a focus on research fraud

Dorothy V. M. Bishop  
Department of Experimental Psychology  
University of Oxford  
OX2 6GG

[dorothy.bishop@psy.ox.ac.uk](mailto:dorothy.bishop@psy.ox.ac.uk)

## Abstract

This report presents analyses of recent annual statements on research integrity by UK universities, summarising the numbers, nature and outcomes of allegations and investigations of research misconduct. Allegations of research misconduct are rare, and focus predominantly on issues such as plagiarism, authorship disputes and failure to obtain ethics approvals. In 117 statements with useable data, there were only 24 allegations of fabrication or falsification involving 12 universities. A complementary analysis categorised PubPeer comments in 2023 for all publications that included a lead UK author. There were 199 commented articles, of which 111 had evidence suggestive of misconduct: 49 of these described fabrication/falsification of data (mostly digitally manipulated images), and 28 described hallmarks of a paper mill. Only nine universities had more than one PubPeer entry compatible with fabrication/falsification, but four of these involved senior researchers with multiple problematic publications; in none of these cases had the institution upheld an allegation of misconduct. It is suggested that those conducting investigations need to be made aware that the research record itself can give *prima facie* evidence of fabrication/falsification. It is recommended that it should not be left to the employing institution to deal with allegations of serious research misconduct by senior staff, that PubPeer could be used proactively in investigations of misconduct, and that research integrity reports should be made openly available to increase confidence in the process.

## Introduction

In a report on a recent Delphi survey, Bishop (2025) noted that there was strong polarisation of views about frequency of serious research misconduct, and adequacy of processes for investigating it. The definition of serious research misconduct encompassed fraud (i.e., fabrication, falsification, plagiarism), and reckless or negligent behaviour with the potential to cause harm. In general, research integrity officers judged that serious research misconduct was not common and current processes for handling allegations, while not perfect, worked reasonably well. In contrast, individuals who undertake independent investigations of research fraud and related issues (often referred to as research "sleuths") regarded rates of serious research misconduct to be high enough to pose a serious threat to science, and concluded that formal processes for investigating and responding to this issue were generally unfit for purpose. Since those who engage research fraud and related misconduct attempt to hide their activities, this topic is hard to investigate. The current article first considers existing evidence for these opposing perspectives, before presenting new empirical data from institutional statements on research integrity from UK institutions, and from PubPeer comments for the same institutions.

## Evidence of research misconduct from surveys

The main source of evidence about rates of research misconduct comes from studies where researchers are asked about their experience of various forms of misconduct. Surveys typically ask respondents to rate how commonly they observe misconduct in colleagues, as well as to give estimates of their own behaviour, because the latter is likely to be an underestimate. The most widely-cited source of evidence of this kind comes from a meta-analysis of survey data by Fanelli (2009), which reported *"A pooled weighted average of 1.97% (N=7, 95%CI: 0.86–4.45) of scientists admitted to have fabricated, falsified or modified data or results at least once ..... In surveys asking about the behaviour of colleagues, admission rates were 14.12% (N=12, 95% CI: 9.91–19.72) for falsification."* Fanelli noted that, given the sensitive nature of the topic, these estimates are likely to be conservative. In a recent preprint (not peer-reviewed), Heathers (2024) attempted to synthesize literature on this topic, and came to a similar conclusion: approximately 1 in 7 scientific papers are either intentionally fabricated or intentionally falsified. However, he also pointed out that the literature is so variable that it does not allow for meta-analysis. Another source is a UK-based study by Williams and Roberts (2020) that used both qualitative (focus group) and quantitative (survey) methods. The survey had 329 responses from research-active academics, and found that around one in five of them reported having fabricated research data. Reports of falsification were much less common. This report was also not peer-reviewed.

Fanelli's review predated the boom in paper mills, commercial outfits that sell authorship of publications for profit. Van Noorden (2023) estimated that 1.5% to 2% of all scientific papers published in 2022 resemble paper-mill works, with biology and medicine being the most affected. Many paper mill products are plagiarised or fabricated. It is not clear how much impact paper mills have on UK universities; most papermill authors are based outside of Europe and the USA (Parker et al., 2024).

## Evidence of research misconduct from RetractionWatch

The Retraction Watch Database is a freely available database which documents all articles that have been retracted. For their Annual Statement of 2024, the UK Committee on Research Integrity (UKCORI) conducted an analysis of all papers in the database with a UK-based author ([UKCORI, 2024](#)). The results appeared reassuring: in the UK, retraction rates were below .05% of all published papers, and comparable to that for other countries in the Organisation for Economic Cooperation and Development. There were 407 articles that were published since 2017 and retracted within two years of publication. Reasons for retraction were "concerns/issues about data" (71 cases), "unreliable results" (54 cases), and "concerns/issues about results". A relatively small number of retractions seems related to paper mill fraud (Abalkina et al., 2025), as evidenced by reasons such as "fake peer review" or "randomly generated content", though it was noted that this may be more of a problem in future. Finally, the UKCORI report noted that retractions do not give definitive evidence of misconduct: authors may ask for their article to be retracted if they find errors in the data or analysis - a sign of research integrity, and some retractions are due to publisher error.

Overall, the analysis of retractions suggested that there is a very low rate of (detected) research misconduct by UK researchers, but there is a serious limitation to this evidence: it can take years for an article to be retracted even when blatant evidence of falsification or fabrication is pointed out to a publisher, and many fraudulent articles remain unretracted (Bik, 2024; Grey et al., 2020; Oransky, 2022).

## Evidence of research misconduct from institutional statements

A further source of evidence that may help throw light on the reasons for polarised views is a set of annual statements that UK universities are required to produce to be compliant with the UK's Concordat to Support Research Integrity (UK Committee on Research Integrity, 2025 - henceforth "the Concordat"). As noted by Chiarelli et al. (2023) in their report on these statements (commissioned by UKCORI): *"The data included in these statements has been undervalued and underutilised. There has been limited analysis of these statements in the past few years, but a thorough and full review of the data within them is expected to offer a more complete understanding of research integrity across the system, provide a baseline from which trends can be monitored and highlight areas where the greatest attention is required."* However, insights from the report are limited because the underlying dataset has been edited to redact quantitative information about misconduct allegations and to anonymise identity of institutions. The authors give the rationale for this approach: *"Throughout this report, and in the supporting data released via Zenodo, we do not name these institutions, except for cases where they are used as part of a good practice case study. This is to avoid the metricisation of our findings, which are meant to paint a picture of research integrity across the UK and not as a means for ranking or comparison."* Furthermore, the focus is on *"narrative discussion rather than quantification, so as to avoid incorrect or inappropriate generalisations that may not reflect the diversity of institutions considered"*.

Data from the Chiarelli et al. report will be presented in more detail below; in general they support the notion that allegations of research misconduct are rare in UK

institutions, and when such allegations occur, they are frequently not upheld. Thus, as with the data from Retraction Watch, estimates of research misconduct from institutional annual statements are considerably lower than estimates obtained by surveys of researchers.

### Evidence of research misconduct from post-publication peer review

Post-publication peer review can provide evidence of potential research misconduct affecting peer-reviewed publications in journals. This will exclude much of the humanities, where books and monographs are a more common form of research output, as well as problematic content in research theses.

PubPeer is a site where anyone can write a review of a published paper, so long as it provides factual support for any statements and does not contain any personal comments about authors. The requirement is that comments contain "facts, logic and verifiable information". Those who report on PubPeer may choose to be anonymous, in which case they are given a botanical name. If their email address is available, authors are informed of comments and encouraged to reply. Although PubPeer was initially set up as a journal club, and comments may be positive or trivial, the majority are critical, and the site has been used to expose major problems in the work of some researchers. It is sometimes represented as a site where malicious allegations are made about researchers, but the moderation policy makes it difficult for bad actors to have a presence (PubPeer, 2024). Comments tend to cluster on certain authors because if a research sleuth discovers an anomaly, they are likely to focus attention on other work by the same author; a single error in an article can always be the result of honest error, but a pattern of errors over a body of work is a strong indication of misconduct. Nevertheless, PubPeer has been criticised as creating a "conundrum" for research integrity investigations, by overwhelming institutions who "*cannot spend every waking hour reviewing PubPeer comments and other potential concerns of research misconduct originating on blogs, social media, and other online sources.*" (Caron, Lye, Bierer, & Barnes, 2024, p. 16).

Another source of evidence is the *For Better Science* website run by Leonid Schneider, which started in 2015 and is now a substantial public record of criticisms of researchers and research studies (Schneider, undated). Although Schneider meticulously documents evidence of research fraud from PubPeer, he goes beyond objective descriptions of errors in research to make accusations of misconduct against researchers, typically drawing on a whole series of publications rather than individual papers, and he includes mockery and personal abuse about researchers while describing suspicious findings in their research. This unprofessional aspect of the website may explain why many institutions adopt antagonistic or dismissive attitudes to *For Better Science*, even though it has been an important source for documenting some high-profile cases of fraud.

The informal impression from PubPeer and *For Better Science* is that serious research misconduct, in the form of fabrication and falsification of data, is common in science, as is implied in the self-report surveys summarised above. However, it remains unclear how far this perception is due to a small number of researchers who engage in repeated

misconduct, or whether it is more widespread. By its very nature, serious research misconduct is usually hidden, making it hard to establish rates of occurrence. In addition, it is unclear how many findings suggestive of serious research misconduct implicate researchers from UK institutions.

## Aims of the current study

The primary aim of the current study is to review the latest annual statements of research integrity from UK institutions, as was done for earlier years in the 2023 report commissioned by UKCORI (Chiarelli et al., 2023). UKCORI has commissioned Research Consulting to do an updated report covering 2023-2024, but this has not yet been published. It seems likely that this will adopt the same approach as the previous report, and so lack transparency about the identity of institutions and not report quantitative data.

By producing a report that includes these variables, the goal is to gain insight into the gulf in perceptions of rates of fraudulent research between research integrity officers and sleuths. The focus was just on the most recent academic year of reporting for which a statement could be found online, typically 2022-2023 or 2023-2024.

To complement this analysis, data from the annual statements were considered in the context of comments on PubPeer for authors from the same institutions, with a particular focus on comments that suggest there has been fraud, i.e., plagiarism, fabrication or falsification of data. For cases where the same research group had several comments suggestive of research fraud, the *For Better Science* website was checked for coverage.

## Methods

### A. Quantitative data from annual statements

As first step, a list of member institutions of [Universities UK](#) was compiled and a web search conducted for the most recent annual statement of research integrity. As noted by Chiarelli et al. (2023), there is no central site for these reports, and so a manual search was required. A search term that combined the name of the institution with the words "research integrity annual" was usually adequate for locating a report, though if this was ineffective, the institution name with "research integrity" generally led to a site that might provide a link to annual reports. It is possible that some reports were missed. The reports used in this analysis were saved in *pdf* format and are available in *Supplement 1*. These reports contain variable amounts of narrative text describing the research integrity procedures adopted by the institution, and topics such as lessons learned from previous experience.

For quantitative data, most reports used a standard template developed by the UK Research Integrity Office (UKRIO, 2022), as shown in Table 1, for reporting numbers of allegations of research misconduct and their outcomes.

*Table 1: UKRIO template for reporting allegations of research misconduct*

Type of allegation	Number of allegations reported to the organisation	Number of formal investigations	Number upheld in part after formal investigation	Number upheld in full after formal investigation
Fabrication				
Falsification				
Plagiarism				
Failure to meet legal, ethical and professional obligations				
Misrepresentation (eg data involvement; interests; qualification; and/or publication history)				
Improper dealing with allegations of misconduct				
Multiple areas of concern (when received in a single allegation)				
Other				
Total				

Instructions for completion are as follows:

*"Please complete the table on the number of formal investigations completed during the period under review (including investigations which completed during this period but started in a previous academic year). Information from ongoing investigations should not be submitted.*

*An organisation's procedure may include an initial, preliminary, or screening stage to determine whether a formal investigation needs to be completed. These allegations should be included in the first column but only those that proceeded past this stage, to formal investigations, should be included in the second column."*

When transferring numbers from reports using this format, the surrounding text was scrutinised and notes added to provide context where possible.

UK institutions cover a massive range in terms of size and research activity, both of which might be expected to affect the number of allegations of research misconduct. To get a rough indicator of both characteristics, two additional variables were

estimated: (a) Size of institution. This was taken from number of students reported in [Wikipedia's List of Universities in the United Kingdom](#) and (b) Number of research outputs. This was obtained by searching the [Dimensions](#) database for research outputs where the Country/Territory of the Research Organisation was United Kingdom for the date range 2020 to 2024.

The database of coded Annual Statements used in the current analysis is available as *Supplementary Material 2*.

## B. PubPeer listings for 2023.

These were provided by PubPeer for articles published in 2023 where any author had an address in the UK. Records were coded for the following variables.

### The identity of the institution.

If more than one UK institution was included, the institution of the first or last authors was used, or, if neither first or last author was from the UK, the first UK institution in the author list was used.

### Institutional contribution

Given that many papers involve authors from several countries/institutions, a measure was devised to reflect the institutional involvement in an article. Each publication was assigned to a single institution, with no double-counting. For the current analysis, the focus is just on articles where the first, last or corresponding author was from a UK institution. If more than one UK institution was represented, then the publication was attributed to the one with most senior author.

### Categorisation of comments in relation to evidence of research misconduct

PubPeer is intended as a platform to report comments about publications, not to report misconduct, and comments alluding to authors' motivations or wrongdoing are moderated out. Nevertheless, the content of comments can indicate problems with published papers that appear to go beyond academic debate, detection of suboptimal methods or honest error.

Accordingly, comments were scrutinised and categorised according to an ad hoc system devised to capture the range of content, with a particular focus on distinguishing between comments suggestive of misconduct and the remainder. There was inevitably a subjective element to this coding, but all records and codes are openly available so that readers can check how justified individual codes are (see *Supplementary Material 3*).

#### Comments not indicative of misconduct

1. Minor errors, such as mislabelling.



2. Detailed review of methods or logic. Some commentators engaged in detailed post-publication peer review of the kind that was originally envisaged for the PubPeer platform. Although these were often critical, this could be seen as robust scientific debate which did not imply misconduct. Where it was difficult to judge without subject-specific expertise, it was assumed that comments on methodological aspects were of this type. Some authors engaged positively with commentators.

3. A serious problem with data that could plausibly reflect honest error where the authors engage to explain this and/or correct it. Errors could reflect duplication of an image within a paper to illustrate groups or conditions that should be different, impossible or implausible data values, or a wrongly reported genetic sequence. In general, this code was used when it was judged that a reputable researcher would not ignore a problem that was this serious.

### **Comments that could indicate misconduct**

4. Serious errors in the data similar to those noted under (3) where the authors fail to give a satisfactory explanation, or obfuscate and/or attack the commentator. Where there was uncertainty about the sufficiency of an author's response, they were given the benefit of the doubt, so code 4 was used only if the author ignored a serious problem or attempted to cover it up. In general, corresponding authors are notified when a PubPeer comment appears, and given the opportunity to respond; however, if the PubPeer record did not show an associated author email, code 3 was used, as it could not be assumed the author was aware of the comment.

5. Undocumented departure from protocol in a registered study.

6. Failure to report ethics approval for a study that requires it.

7. Undeclared conflict of interest.

8. Authorship/affiliation issues. This included guest authorship, or misleading affiliation or email. This code excluded articles with the hallmarks of a paper mill (see below).

### **Comments strongly suggestive of research fraud**

9. Evidence of plagiarism or self-plagiarism

10. Evidence that the article comes from a paper mill. Paper mill outputs typically involve elements of fabrication or falsification but were categorised separately here, because they have the distinctive feature that researchers have purchased authorship. They can also be identified by other characteristics, including evidence from the Problematic Paper Screener (Cabanac, Labbé, & Magazinov, 2022), which notes "tortured phrases", citation of retracted, questionable or otherwise unreliable sources, and indicators of compromised editorial or peer review processes. Another indicator is numerous co-authors from many different countries in a context where this is hard to explain, or a co-author with a strong track record of paper mill publications. In general, there is no one definitive indicator of paper mill products. This code would not be

applied if there was just a single instance of a tortured phrase or inappropriate citation in an article; rather the judgement is made from the presence of more than one "red flag" such as these. Finally, some paper mill products can be identified from a retraction notice; the publisher Hindawi retracted thousands of papers after its journals became infested with paper mill products (Bik, 2023).

11. Evidence of data fabrication or falsification. Fabrication (making up data) and falsification (altering existing data) can be hard to distinguish and so are treated together here. Some PubPeer comments included screenshots that show manipulations of data that are difficult to explain other than as a deliberate attempt to mislead. The main category here is digital manipulation of images. This goes beyond the kind of image duplication described in (3) to include overwriting sections of an image, or rotating, stretching or otherwise manipulating an image to hide the fact it is a duplication.

Comments that did not belong in any of the categories 1-11 were coded as Other (12).

### Author response

A code was added to indicate whether the author responded on PubPeer to a comment.

### Journal action

This was coded as none, corrected, expression of concern (EOC), or retraction.

### Cross-check against other sources

Where the same author names recurred across multiple comments, a web search was conducted to discover whether there had been significant PubPeer comments associated with their name in other years, and whether a web search yielded further information.

### Analysis plan

The analysis is descriptive, and no inferential statistics are used. Data from the Annual statements were tabulated in a format similar to that used by Chiarelli et al. (2023) for their analysis of earlier years. The Chiarelli et al. data for 2019-2020 and 2020-2021 are displayed alongside the current analysis of most recent reports. The Chiarelli et al. data for 2021-2022 are excluded because the authors reported that there was a relatively small number of reports because several institutions were in the process of completing their statement at the time of their data collection. In contrast to Chiarelli et al, the current analysis included tabulation of allegations involving fabrication and falsification, indicating which institutions had such cases and how the allegations had been handled.

For PubPeer comments, descriptive data for different categories of comment are shown, with a more in-depth analysis of institutions with two or more instances of any category suggestive of research misconduct.

## Results

### A. Quantitative data from annual statements

Table 2 shows the numbers of institutions providing an annual statement; data from earlier years from Chiarelli et al. (2023) are provided for comparison. Data from 2021-2022 from Chiarelli et al. are excluded. Note that the current analysis takes data from the latest available statement, which in 56 cases was from 2022-2023 and in 85 cases was from 2023-2024. The data are coded in the same way as in the report by Chiarelli et al. (2023), using the categories from the reporting template shown in Table 1.

*Table 2: Numbers of allegations in annual statements compared with prior years*

Statistic	2019-2020*	2020-2021*	2022-2023 or 2023-2024**
N institutions in scope	134	134	141
N (%) providing an annual statement	97 (72%)	104 (78%)	117 (83%)
N providing useable quantitative data***			109
Statements reporting at least one misconduct allegation (% of those with useable data)	60 (62%)	60 (58%)	60 (55%)
Statements reporting at least one misconduct investigation (% of those with useable data)	50 (52%)	53 (51%)	45 (41%)
Statements reporting at least one allegation upheld in full (% of those with useable data)	22 (23%)	23 (22%)	27 (25%)
Statements reporting at least one allegation upheld in part (% of those with useable data)	9 (9%)	6 (6%)	12 (11%)

\*Data from table on p 13, Chiarelli et al. (2023).

\*\*Analysis based on most recent annual statement: 56 were from 2023 and 85 from 2024.

\*\*\* Chiarelli et al. stated that 98% of statements in their corpus provided data on misconduct allegations, investigations and outcomes

As anticipated, institutions that did not produce a statement tended to be smaller and less research-intensive than those that did: for the 24 institutions producing no statement, the 25%, 50% and 75% quantiles for size (Wikipedia, N students) were 6,765, 10,920 and 16,725, and for the 117 with a statement the corresponding figures were 12,052, 17,540, 25,910. Similarly, for the total N publications indexed in Dimensions from 2020-2024, the 25%, 50% and 75% quantiles for those with no statement were 0, 1,324 and 4,571, whereas the corresponding quantiles for those with a statement were 1,498, 4,587, 12,014. All 26 institutions with 15K or more publications in 2020-2024 produced a statement.

In many respects, the data from the current analysis are similar to those from previous years, with 55% of institutions reporting one or more allegations of research

misconduct, and the remaining 45% reporting no cases. The percentage of allegations that proceeded to formal investigation appears to have declined somewhat from 2019-2021, but the outcomes of investigations were similar: 25% of institutions had at least one allegation that was upheld in full and 11% institutions had at least one allegation that was upheld in part. Overall, in one year, just over half all institutions received allegations of research misconduct, and around two thirds of allegations were not upheld.

For the current analysis some statements could not be included in computations of numbers of allegations etc., because they either omitted key information or appeared to misinterpret the instructions for numeric data. There was also evidence of different interpretations of what needed to be included in the table, and it was sometimes unclear how to interpret the distinction between an allegation that was not acted on and one that proceeded to investigation. A few institutions included cases of misconduct in undergraduate documents, whereas most treated these under the separate category of "academic misconduct". Available data were interpreted taking into account any accompanying text, but the quantitative data needs to be interpreted cautiously, given the uncertainties in the underlying reports. Explanatory notes are provided in the underlying data table to show transparently what interpretive decisions were made (see *Supplement 2*).

There were 91 institutions with data both on number of publications in the Dimensions database, and with number of allegations. To get a rough estimate of the proportion of publications associated with an allegation in one year, the total number of allegations for this subset of institutions (221), was divided by the total number of publications in 2024 (263,122). Thus the proportion of allegations per publication was .00084, or less than one allegation per thousand publications. Not all allegations concern published papers, and so this is if anything an overestimate. This figure contrasts strikingly with estimates of falsification and fabrication from surveys summarised above, which suggest that as many as 14% researchers may produce fraudulent work.

Amount of research activity was again a factor affecting these results; as one might anticipate, the larger the number of publications associated with an institution, the larger the number of allegations. For those had no allegations, the 25%, 50% and 75% quantiles for N publications 2020-2024 were 0, 1,775, and 4,774, whereas for those with allegations, the corresponding figures were 4,196, 8,586, 28,454. This would be expected even if rates of misconduct were constant across institutions, because the more publications, the greater the likelihood that a problem will be found in a publication.

Table 3 turns the focus from the number of institutions with allegations to the raw numbers of misconduct allegations summed across institutions. The Table shows numbers that proceed to formal investigation and/or which were ultimately upheld.

*Table 3: Fate of misconduct allegations*

Variable	2019-2020*	2020-2021*	2022-2023 or 2023-2024
N allegations (totalled across institutions)	283	277	226
N proceeding to formal investigation (% of allegations)	183 (65%)	154 (56%)	92 (41%)
N where allegation upheld in part or full (% of investigations)	58 (32%)	95 (61%)	55 (60%)

\*From Chiarelli et al. (2023) Table on p. 14.

Compared to the previous years considered by Chiarelli et al. (2023) the number of allegations in the statements analysed for 2022-2023/2023-2024 is lower, and when an investigation is made it is less likely to proceed to formal investigation.

In the current dataset, 60% of misconduct allegations that proceeded to a full investigation, were upheld - a figure comparable to that for 2020-2021. These numbers are based on aggregating small numbers across many universities with a very diverse range of types of misconduct allegation, ranging from authorship dispute to fabrication and falsification of data. As noted by Chiarelli et al (2023), another factor making it difficult to interpret results is inclusion of the same case in more than one report if an investigation takes over a year, or if more than one institution is involved. Nevertheless, we can conclude that most allegations of research misconduct are resolved informally without going to formal investigation. In the 2022-2023/2023-2024 data, only around one in four allegations was ultimately both investigated and upheld.

Some allegations can readily be judged on the basis of documentary evidence, such as ethics approval documents, or duplication of content across documents that confirms plagiarism; others require judgement of adherence to rules such as authorship guidelines. Allegations of fabrication or falsification of data typically can be evaluated using evidence from published articles, but may be difficult to judge for those without subject-specific expertise.

Table 4 shows the overall distribution of categories of misconduct found in allegations, again aggregated across UK institutions. Data are from the latest report (2022-2023 or 2023-2024), for the subset of statements that coded the type of misconduct.

*Table 4. Categories of misconduct in most recent annual statements*

Variable	Total allegations	Institutions reporting at least one case
N allegations that could be categorised (totalled across institutions)	226	
Plagiarism	74	30
Failure to meet legal, ethical and professional obligations	72	33
Fabrication or falsification	24	12
Other (including multiple types)	47	26

This distribution of categories of misconduct is broadly similar to that reported by Chiarelli et al. (2023), who found across all in-scope years there were 101 reports giving a breakdown of categories. These included 203 allegations of plagiarism, 90 cases of failure to meet legal, ethical and professional obligations, and 63 cases of misrepresentation. The categories of fabrication and falsification were not mentioned, presumably because they were so uncommon.

For the current dataset, we took a more detailed look at the fate of cases of fabrication and falsification, because this type of misconduct is among the most serious.

*Table 5. Details of cases of fabrication/falsification*

Institution	N cases	N upheld	notes
University College London, 2023-2024	4	0	no further details except some "closed off at the initial assessment stage" and others "closed off at the screening stage"
University of Oxford, 2023-2024	3	0	all dismissed after preliminary review - no evidence of misconduct
Imperial College, 2023-2024	1	0	(2 further cases listed as ongoing; not considered in this analysis) no further information provided
University of Manchester, 2023-2024	1	1	One case partially upheld
University of Nottingham, 2023-2024	1	1	One PGR student was investigated for academic misconduct, with an allegation upheld. This involved fabrication of reference and false authorship in use of AI software
University of Sheffield, 2023-2024	1	0	no formal investigation
University of Southampton, 2023-2024	1	0	no formal investigation
Queen Mary University of London, 2023-2024	5	0	Each case described with case number.

			<p>8-9: Two separate formal complaints from the same individual were judged to have insufficient evidence of misconduct; also one complaint was "comprised entirely of links to comments posted on PubPeer, ... this did not amount to an actionable complaint."</p> <p>10: Allegation of fabrication and uncorrected errors in figures. This case had previously been considered and actions had been recommended, but the journal had not acted.</p> <p>11. Former PhD student alleged falsification in a published journal article. Expert opinion found errors in some of the statistical equations presented, but not sufficient to alter the fundamental conclusions of the paper.</p> <p>12. Complaint of citation manipulation did not provide substantiation beyond including links to posts on PubPeer - not provided sufficient evidence for a formal investigation.</p>
Cardiff University	2	0	<p>(Case 13 is not concluded so omitted here).</p> <p>1. A Journal noted comments relating to image manipulation on Pubpeer. This had not been raised as an allegation of ARM with the University, but Head of School carried out a proactive review of the Pubpeer comments. Since the paper involved collaboration between Cardiff and another university, the corresponding author was tasked with reviewing and responding to the specific issues raised on PubPeer with oversight from the relevant university partners.</p> <p>2. Allegation of image manipulation was concluded after screening panel accepted respondent's statement that this was genuine error, and not intentional and did not significantly impact the overall findings of the paper.</p>
Swansea University, 2023-2024	1	0	no details
De Montfort University, 2022-2023	1	1	A research student admitted to the fabrication of results and the Academic Offense Panel imposed a penalty.
Middlesex University, 2023-2024	2	0	no details

Overall, 3 of 24 allegations of fabrication or falsification were upheld. Two of these involved postgraduate students; no information is provided about the other one.

## B. PubPeer comments for 2023

Table 6 shows the numbers of PubPeer comments for each type of classification across all UK universities.

*Table 6: PubPeer comments classification, by N comments and N universities*

Categories of comments	Total N papers with comments	N papers with comments where main author based at UK university	N UK universities with 2 or more papers in this category (only those where main author at that university)
1.Minor	20	11	1
2.Critique	65	41	6
3.Data_error	44	30	7
4.Data_error_questionable	44	24	6
5.Protocol_error	3	3	0
6.Ethics_breach	5	2	0
7.COI	6	1	0
8.Authorship	3	1	0
9.Plagiarism	9	4	0
10.Papermill	62	27	3
11.Fabrication_falsification	74	49	9
12.Other	13	6	1

Table 6 illustrates one problem of using individual publications as evidence for a misconduct allegation: many articles involve collaboration across institutions, which can make it difficult to determine responsibility for conducting an investigation. The column with Total N papers with comments covers many publications where the principal author was outside the UK. The focus of the current analysis is on the next column, showing articles where the first or last author is at a UK institution. Some of the categories that were relatively common in annual statements barely feature in Table 6; PubPeer comments were rare for protocol errors, ethics breaches, conflicts of interest, authorship disputes, and plagiarism. Furthermore, as shown in the last column, they are not evenly distributed across institutions: note, for instance, that only three institutions had two or more comments suggestive of papermill activity and only nine institutions had two or more comments suggestive of fabrication or falsification.

In this analysis of PubPeer comments for 2023, there were two cases when a research integrity officer - from the University of Dundee - responded to a critique of work on PubPeer (<https://pubpeer.com/publications/36FD011475B9977649B1868A5E9BD9>; <https://pubpeer.com/publications/72FBFBA261EA1B463B204D44173F01>). Apart from this isolated instance, however, there was no indication that institutions engaged with PubPeer as a means of monitoring research outputs.



Table 7 looks in more detail at PubPeer comments that were in categories that suggested research misconduct. The Table includes only the 14 universities where there were two or more papers with a comment in a given category. Data are shown using a more lax criterion, where any university with a comment coded as 4-11, in *Supplement 4*. There were 42 universities meeting that criterion.

The clustering of comments in Table 7 that characterises this table is not accidental. As noted above, a sleuth who finds evidence of falsification or fabrication of data will be motivated to check other work by the author(s) for confirmatory evidence. A single article with a manipulated image is unlikely to be taken as sufficient grounds for a research integrity investigation, especially given that most publications have multiple authors. Unless there is an author credit statement, it may be unclear who was responsible for the manipulation. Scrutiny of a wider range of outputs by the same authors will usually provide an indication of whether there is a persistent pattern of fabrication/falsification, and which authors are associated with it.

*Table 7: Characteristics of comments that implied misconduct, grouped by institution*

Institution	Category	N	Notes
Bristol	Data error questionable	2	(1-2) All relate to one research group: This case featured in <a href="#">For Better Science</a> .
Bristol	Fabrication falsification	10	(1-10) Two articles have Expression of Concern. All ten relate to one group, which had featured in <a href="#">For Better Science</a> .
Cambridge	Fabrication falsification	2	(1) PI has 2 previous PubPeer comments re image splicing; (2) EOC for this paper; Author has no other PubPeer
Cardiff	Data error questionable	3	(1-3) All cases where there was no response to queries re figure duplications. These all relate to one research centre that was featured on <a href="#">For Better Science</a> in <a href="#">2018</a> : and again in <a href="#">2023</a> .
Cardiff	Fabrication falsification	2	(1) This comment relates to same centre as previous row. (2) All authors from Pakistan except last (senior) author from Cardiff Pharmacy. He has PubPeer comments on other papers reporting image issues.

Durham	Papermill	2	<p>(1) Out-of-scope article in special issue; author has dual affiliation Durham Business School and Chinese university; no other obvious PubPeer (though name disambiguation difficult); Google search on this name with "Durham" yields no hits, suggesting this may have been a student who has now left.</p> <p>(2) Retracted article from tanu.pro paper mill; whole special issue retracted</p>
Imperial	Data error questionable	2	<p>(1-2) Both comments relate to articles by the same Professor in collaborative articles with China; he has other papers commented on in PubPeer and is mentioned in <a href="#">For Better Science</a>.</p>
KCL	Data error questionable	3	<p>(1) Has correction with remarkable number of figure changes; Senior author has co-authored some other papers with PubPeer queries re data/figures but not as corresponding author.</p> <p>(2) Senior author co-authored one other paper years ago with queries that were resolved.</p> <p>(3) 1st author was at KCL, now Japan; last author is Prof at KCL with other PubPeer comments on articles with image queries.</p>
Leeds	Fabrication falsification	2	<p>(1) Retracted for unresolved figure irregularities; Neither 1st nor last author has any other PubPeer comments;</p> <p>(2) Journal correction for figure anomalies, though these involve more than duplication; Last author has no other PubPeer comments</p>
Leicester	Fabrication falsification	2	<p>(1) Senior author has other PubPeer comments for image manipulation and COI; also featured in <a href="#">For Better Science</a>.</p> <p>(2) Senior author has co-authored other papers with PubPeer image concerns, but not as corresponding author.</p>
Lincoln	Fabrication falsification	2	<p>(1) Last author Professor of Nanoscience from Lincoln; some engagement on PubPeer but not convincing; article retracted; He is featured in <a href="#">For Better Science</a>.</p>

			(2) Last author Associate Professor of Pharmacology and Therapeutics; features in <a href="#">For Better Science</a> .
Manchester	Fabrication falsification	3	(1-2) Same corresponding author: Senior Lecturer; although author engaged on PubPeer did not convincingly address issues; (3) last author dual appointment with Manchester Biomedical Research Centre. Features in <a href="#">For Better Science</a> .
Portsmouth	Papermill	2	(1) Retracted. Last author joint affiliation with Portsmouth and Saudi Arabia; (2) Retracted; last author Emeritus Professor of Pharmacognosy
QMUL	Data error questionable	3	(1-3) All with same senior author, featured in <a href="#">For Better Science</a> .
QMUL	Fabrication falsification	15	(1-8) All with same senior author as for previous row. Two retracted, one EOC. Featured here: <a href="#">For Better Science</a> .  (9-14) One journal correction. Author featured in <a href="#">For Better Science</a> . (15) authors from Argentina except for last author, Emeritus Professor of Biochemistry, now deceased.
Salford	Papermill	4	(1-4) All co-authored by Professor of Fluid Mechanics, Propulsion and Nanomechanics Research
UCL	Data error questionable	2	(1) Last author Director of The KCL Centre for Cell and Gene Therapy; featured in <a href="#">For Better Science</a> regarding another article (2) last author is Associate Professor; has one prior PubPeer comment where he engaged positively
UCL	Fabrication falsification	3	(1) Last (corresponding) author has honorary position at Cancer Institute. He has coauthored several articles flagged on PubPeer for image manipulation;  (2-3) First author was research fellow at UCL, now at Swansea. He has co-authored one other paper noted on PubPeer for image manipulation concerns.

The clusters of comments seen for Bristol, Cardiff and QMUL (x 2) appear to provide strong enough evidence for an allegation of research misconduct to be taken seriously, particularly when one bears in mind that this table is confined to PubPeer comments from 2023, so additional comments may be found by searching other years.

Nevertheless, it seems that in none of these cases was a case upheld. Nearly all the comments for QMUL related to one of two research groups. In the annual statement, the QMUL research integrity officer noted two allegations that were judged to have insufficient evidence of misconduct; one complaint was "comprised entirely of links to comments posted on PubPeer, ... this did not amount to an actionable complaint."

## Discussion

The quantitative analysis of the most recent annual statements by UK universities obtained results broadly similar to those reported by Chiarelli et al (2023) for previous years. In brief:

- Allegations of research misconduct are rare, particularly in smaller universities with a primary focus on teaching. Across the whole set of universities, there was less than one allegation per thousand publications.
- The most common types of research misconduct that feature in annual statements are plagiarism and failure to meet legal, ethical and professional obligations. There is insufficient data to be certain, but the impression is that these types of misconduct mostly affect graduate students rather than academic staff.
- Research fraud other than plagiarism, i.e. fabrication/falsification of data, is rarely reported: there were 24 cases from 12 universities, only three of which were upheld.

This analysis was complemented by an analysis of comments on publications on the post-publication peer review website PubPeer. The focus here is on comments relating to articles where the main author(s) are from a UK institution.

- As shown in Table 6, the number of publications from UK institutions attracting comments is relatively small. Even if every PubPeer comment were investigated, there seems no danger of UK Research Integrity Officers spending "*every waking hour reviewing PubPeer comments*" (Caron, Lye, Bierer, & Barnes, 2024, p. 16).
- Around 20% of comments concerned discussion of aspects of methods, logic or results that appear to correspond to regular academic debate. A further 20% are either minor comments, or are consistent with "honest error", where the author engages with explanation or correction.
- Other comments reported more serious issues that appear indicative of research misconduct: 14% of comments related to publications that seem to be associated with paper mills; 25% of comments provided evidence of data

fabrication/falsification, and a further 12% reported serious errors in articles that were not adequately addressed by authors. Only 2% of comments related to plagiarism.

- There were 14 universities with two or more comments that suggested serious research misconduct, and in four cases there were several such comments focused on specific research groups. Although these cases had received critical commentary on blogs, and some of the publications had had expressions of concern, no cases of research misconduct relating to these had been upheld by institutions, despite what appeared to be *prima facie* evidence of fabrication/falsification of data.

Taken together, both annual statements and PubPeer reports find a low frequency of reports of problems. This might seem reassuring, were it not for the discrepancy with estimates of research misconduct from surveys, which imply that many cases of research misconduct go undetected or unreported.

It is also clear that allegations of fabrication and falsification are an uncommon reason for institutional research misconduct investigations, which focus predominantly on plagiarism, authorship disputes and breach of professional standards, including ethics approvals.

The greatest concern to emerge from this analysis is that in the rare cases where serious allegations of fabrication/falsification are considered, they are typically dismissed by institutions. Of course, it cannot be assumed that all such allegations are well-founded, but the evidence from PubPeer is typically unarguable in cases where one can see that images have been manipulated. While there may be debate about who is responsible for the manipulation, repeated comments regarding data fabrication/falsification in different publications from the same research group should be taken seriously.

## Conclusions and recommendations

We started by considering the mismatch in views of research integrity investigations by research integrity officers and sleuths that was apparent in the FAIRS survey by Bishop (2025). The current analyses give some reassurance that, consistent with the view of research integrity officers, the numbers of allegations of serious research misconduct in UK universities are low. Although some institutions are still reporting in an ad hoc fashion, it appears that the Concordat is helping UK institutions to follow common standards (cf. the situation in the USA: Gunsalus, Marcus, & Oransky, 2018). However, consistent with the views of sleuths, it appears that when serious cases of research fraud occur, they seldom lead to the institution concluding that there was research misconduct, even where there is strong evidence of fabrication or falsification in more than one publication.

One point to emerge from the analysis of annual statements is that many research integrity officers have no experience with data falsification or fabrication, and may lack

confidence in determining what might be plausible honest error. It is not uncommon for a PubPeer commentator to briefly indicate areas of figures with overlap using coloured rectangles. A naive reader may not recognise when this is indicative of digital manipulation and hence fail to appreciate it is evidence of fabrication/falsification of data (Rossner & Yamada, 2004). Those making allegations need to spell out the issues that have been identified at a level of detail that can be understood by a non-expert, so that a defence of "honest error" is not accepted when it is not plausible.

A particular issue frequently alluded to in the FAIRS survey (Bishop, 2025) as well as in allegations on the *For Better Science* website, is institutional conflict of interest. Some allegations of research fraud involve senior figures who hold important institutional positions and bring in large amounts of grant income. While it may be reasonable for institutions to investigate matters such as authorship disputes or plagiarism in postgraduate theses, it is questionable whether they can realistically be expected to be objective when investigating allegations against their own senior researchers. As noted by Redman (2023) the inherent conflict of an institution investigating itself and its faculty is *"a flaw that should require independent outside review"* (p 115-6). It would take pressure off research integrity officers and give greater confidence in the process if allegations of fraud in the form of falsification and fabrication of data were handled by a body independent of the institution. Some countries have created such bodies for exactly this situation, but in the UK this is not the case (National Health and Medical Research Council, 2022).

A third issue concerns whether institutional research integrity officers should be more proactive in tackling research integrity matters, or continue to adopt an exclusively responsive mode where an investigation is initiated only when an allegation is received in an appropriate format. Although PubPeer is far from comprehensive in its coverage, it could help institutions get ahead of integrity issues if they were able to scan PubPeer for comments affecting their researchers and encourage authors to resolve these if possible.

A final point concerns transparency of research integrity investigations. The prior UKCORI report (Chiarelli et al., 2023) stated that quantitative data were omitted to *"avoid metricisation"*, and identity of institutions was masked to avoid *"incorrect or inappropriate generalisations that may not reflect the diversity of institutions considered"*. However, the wide variety in allegations of research misconduct and responses to them is one reason why it can be useful to look at institutions in more granular detail. The aggregate figures alone are not informative for guiding future behaviour. Here the analysis documented individual cases, which may help one identify where there is good practice and where there are problems.

Oransky and Marcus (2025) took a step further and made a cogent case that reports of misconduct investigations should be made public, as has been done in the past by the US Office of Research Integrity. Currently, investigations referred to in annual statements by UK universities are anonymised. This means that where an allegation of misconduct is upheld, it is easy for the researcher in question to relocate and continue fraudulent activities elsewhere. For instance, Retraction Watch (2025) reported that an

investigation by Manchester University had confirmed that one of their former students had been running a paper mill, but it is unclear whether this case was included in their annual statement for 2023-4, and it is evident that the student went on to another university. It is at least as important to be fully transparent when an investigation concludes there is no misconduct: if the process is shrouded in secrecy, then this creates suspicion of a cover-up, which does not do any favours to the researchers accused of misconduct. This is illustrated in a case from Queen Mary University where statements of "no misconduct" about one researcher have been publicly challenged (Sinha, 2023). Virtually every UK university has a section on their website that echoes sentiments similar to those stated here by UCL *"All staff and students at UCL are expected to promote and maintain a culture of honesty, openness and responsibility, enabling all research to be conducted with integrity. Ensuring the highest standards of integrity in all aspects of our research activities allows others to have trust and confidence in the methods used and the findings."* Greater transparency in reporting of investigations could help garner such trust and confidence.

## References

- Abalkina, A., Aquarius, R., Bik, E., Bimler, D., Bishop, D., Byrne, J., Cabanac, G., Day, A., Labbé, C., & Wise, N. (2025). 'Stamp out paper mills'—Science sleuths on how to fight fake research. *Nature*, 637(8048), 1047–1050. <https://doi.org/10.1038/d41586-025-00212-1>
- Bik, E. M., Casadevall, A., & Fang, F. C. (2016). The prevalence of inappropriate image duplication in biomedical research publications. *mBio*, 7(3), e00809-16, /mbio/7/3/e00809-16.atom. <https://doi.org/10.1128/mBio.00809-16>
- Bik, E. (2023, August 11). Hindawi's mass retraction of "Special Issues" papers. *Science Integrity Digest*. <https://scienceintegritydigest.com/2023/08/10/hindawis-mass-retraction-of-special-issues-papers/>
- Bik, E. M. (2024). Finding the truth in science. *Nature Reviews Cancer*, 24(8), 516–517. <https://doi.org/10.1038/s41568-024-00702-w>
- Bishop, D. V. M. (2025). *FAIRS Delphi survey: Overview of results*. <https://osf.io/kjsn4>
- Cabanac, G., Labbé, C., & Magazinov, A. (2022). *The 'Problematic Paper Screener' automatically selects suspect publications for post-publication (re)assessment* (arXiv:2210.04895). arXiv. <https://doi.org/10.48550/arXiv.2210.04895>
- Caron, M. M., Lye, C. T., Bierer, B. E., & Barnes, M. (2024). The PubPeer conundrum: Administrative challenges in research misconduct proceedings. *Accountability in Research*, 1–19. <https://doi.org/10.1080/08989621.2024.2390007>
- Chiarelli, A., Carter, I., Palmer, F., Thomas, B., Higgs, J., & Johnson, R. (2023). *Review of annual statements on research integrity*. Zenodo. <https://doi.org/10.5281/zenodo.8014377>
- Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLOS One*, 4(5), e5738. <https://doi.org/10.1371/journal.pone.0005738>
- Grey, A., Avenell, A., Gamble, G., & Bolland, M. (2020). Assessing and raising concerns about duplicate publication, authorship transgressions and data errors in a body of preclinical research. *Science and Engineering Ethics*, 26(4), 2069–2096. <https://doi.org/10.1007/s11948-019-00152-w>
- Gunsalus, C. K., Marcus, A. R., & Oransky, I. (2018). Institutional research misconduct reports need more credibility. *JAMA*, 319(13), 1315–1316. <https://doi.org/10.1001/jama.2018.0358>
- Heathers, J. (2024). *Approximately 1 in 7 scientific papers are fake*. <https://osf.io/s4gce>



- House of Commons: Science and Technology Select Committee. (2018). *Research Integrity*.  
<https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/350/35002.htm>
- National Health and Medical Research Council. (2022). *Compilation of information about research integrity arrangements outside Australia* (OECD Science, Technology and Industry Policy Papers 130; OECD Science, Technology and Industry Policy Papers, Vol. 130). <https://doi.org/10.1787/1c416f43-en>
- Oransky, I. (2022). Retractions are increasing, but not enough. *Nature*, 608(7921), 9–9.  
<https://doi.org/10.1038/d41586-022-02071-6>
- Oransky, I., & Marcus, A. (2025). Why universities should make misconduct reports public. *Journal of Law, Medicine & Ethics*, 1–5. <https://doi.org/10.1017/jme.2025.39>
- Parker, L., Boughton, S., Bero, L., & Byrne, J. A. (2024). Paper mill challenges: Past, present, and future. *Journal of Clinical Epidemiology*, 176, 111549.  
<https://doi.org/10.1016/j.jclinepi.2024.111549>
- Redman, B. (2023). *Reconstructing Research Integrity: Beyond Denial*. Springer.
- PubPeer. (2024). *Commenting Guidelines*. <https://blog.pubpeer.com/static/faq#12b>
- Retraction Watch. (2025, January 30). Former student was running a paper mill, says University of Manchester. *Retraction Watch*.  
<https://retractionwatch.com/2025/01/30/former-student-was-running-a-paper-mill-says-university-of-manchester/>
- Rossner, M., & Yamada, K. M. (2004). What's in a picture? The temptation of image manipulation. *Journal of Cell Biology*, 166(1), 11–15.  
<https://doi.org/10.1083/jcb.200406019>
- Schneider, L. (n.d.). *For Better Science*. Retrieved 17 May 2025, from <https://For Better Science.com/>
- Sinha, G. (2023, November 15). *University says it found no misconduct in anti-inflammatory research. Critics are unconvinced*.  
<https://www.science.org/content/article/university-found-no-misconduct-anti-inflammatory-research-critics-unconvinced>
- UK Committee on Research Integrity. (2024). *Annual statement of the UK Committee on Research Integrity 2024*. <https://ukcori.org/wp-content/uploads/2024/07/UK-Committee-on-Research-Integrity-Annual-Statement-2024-1.pdf>
- UK Committee on Research Integrity. (2025). *The Concordat to Support Research integrity, 3rd edition*. <https://ukcori.org/wp-content/uploads/2025/04/The-Concordat-to-Support-Research-Integrity-2025.pdf>

UKRIO. (2022, November 15). *Annual statement on research integrity template—UK Research Integrity Office*. <https://ukrio.org/news/annual-statement-template/>

Van Noorden, R. (2023). How big is science's fake-paper problem? *Nature*, 623(7987), 466–467. <https://doi.org/10.1038/d41586-023-03464-x>

Williams, J., & Roberts, D. (2020). Academic Integrity: Exploring Tensions Between Perception and Practice in the Contemporary University. Society for Research into Higher Education. <https://srhe.ac.uk/wp-content/uploads/2020/03/WILLIAMSJoannaROBERTSDavid.pdf>.

## Acknowledgements

My thanks to Csaba Szabo, René Aquarius and George Corby for helpful comments on the draft of this report.

## Funding

This work received no external funding.

## Declaration of interests

The author reports no conflict of interests.