The Prevalence of Inappropriate Image Duplication in Biomedical Research Publications

Elisabeth M. Bik[1#], Arturo Casadevall*[2,3], Ferric C. Fang*[4]

[1]Department of Medicine, Division of Infectious Diseases, Stanford School of Medicine, Stanford, California, USA

[2]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

[3]Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

[4]Departments of Laboratory Medicine and Microbiology, University of Washington School of Medicine, Seattle, Washington, USA

Running title: **Image duplication prevalence in biomedical papers**

#Address correspondence to: Elisabeth M. Bik, eliesbik@stanford.edu

VAHCS Palo Alto, 3801 Miranda Ave, Palo Alto, CA 94304. Telephone: +1 650-493-5000, extension 63163

*AC and FCF contributed equally to this work

Abstract word count: 154

Text word count (excluding references and figure legends): 4064

ABSTRACT

Inaccurate data in scientific papers can result from honest error or intentional falsification. This study attempted to determine the percentage of published papers containing inappropriate image duplication, a specific type of inaccurate data. The images from a total of 20,621 papers in 40 scientific journals from 1995-2014 were visually screened. Overall, 3.8% of published papers contained problematic figures, with at least half exhibiting features suggestive of deliberate manipulation. The prevalence of papers with problematic images rose markedly during the past decade. Additional papers written by authors of papers with problematic images had an increased likelihood of containing problematic images as well. As this analysis focused only on one type of data, it is likely that the actual prevalence of inaccurate data in the published literature is higher. The marked variation in the frequency of problematic images among journals suggest that journal practices, such as pre-publication image screening, influence the quality of the scientific literature.

Keywords: Research misconduct, ethics in science, biomedical research, peer review

IMPORTANCE

The scientific literature is cumulative, and the reproducibility of biomedical research is a topic of increasing concern. Inaccurate data, whether resulting from honest error or intentional misconduct, can contribute to research irreproducibility, but the prevalence of such data is unknown. Here, we provide the first estimate of the percentage of a specific type of inaccurate data, inappropriate image duplication, in published biomedical research papers. Approximately 1 of every 25 papers was found to contain some form of problematic image, with a substantial increase in prevalence since 2003. Current standards appear insufficient to prevent flawed papers from being published. Greater efforts are needed to ensure the reliability and integrity of the research literature.

51 **INTRODUCTION**

52

53 Inaccuracies in scientific papers have many causes. Some result from honest mistakes,
54 such as incorrect calculations, use of the wrong reagent, or improper methodology (1). Others
55 are intentional and constitute research misconduct, including situations in which data are
56 altered, omitted, manufactured or misrepresented in a way that fits a desired outcome. The
57 prevalence of honest error and misconduct in the scientific literature is unknown. One review
58 estimated the overall frequency of serious research misconduct, including plagiarism, to be 1%
59 (2). A meta-analysis by Fanelli, combining the results of 18 published surveys, found that 1.9%
60 of researchers have admitted to modification, falsification or fabrication of data (3).

61 There is also little firm information on temporal trends regarding the prevalence of error
62 and misconduct. Research error and misconduct have probably always existed. Even scientific
63 luminaries such as Darwin, Mendel, and Pasteur have been accused of manipulating or
64 misreporting their data (4, 5). However, the perception of error and misconduct in science has
65 been recently magnified by high profile cases and a sharp rise in the number of retracted
66 manuscripts (6). In recent years, retractions have increased at a rate that is disproportionately
67 greater than the growth of the scientific literature (7). Although this could be interpreted as an
68 increase in problematic papers, the actual causes may be more complex and could include a
69 greater inclination by journals and authors to retract flawed work (7). Retractions are a poor
70 indicator of error because most retractions result from misconduct (8), and many erroneous
71 studies are never retracted (1). In fact, only a very small fraction of the scientific literature has

been retracted. As of April 2016, the PubMed bibliographic database listed 8,735 retracted publications among more than 25 million articles (0.035%).

Concerns about misconduct have been accompanied by increasing concerns about the reproducibility of the scientific literature. An analysis of 53 landmark papers in oncology reported that only 6 could be reproduced (9), and other pharmaceutical industry scientists have also reported low rates of reproducibility of published findings, which in some cases led to the termination of drug development projects (10). In the field of psychology, less than half of experimental and correlational studies are reportedly reproducible (11). Inaccurate data can result in societal injury. For example, a now-retracted study associating measles vaccination with autism continues to resonate and may be contributing to low vaccination rates (12). Corrosion of the literature, whether by error or misconduct, may also impede the progress of science and medicine. For example, false leads may be contributing to increasing disparities between scientific investment and measurable outcomes, such as the discovery of new pharmacological agents (13).

In this study we sought to estimate the prevalence of a specific type of inaccurate data that can be readily observed in the published literature, namely inappropriate image duplication. The results demonstrate that problematic images are disturbingly common in the biomedical literature and may be found in approximately 1 out of every 25 published articles containing photographic image data, in particular Western blots.

RESULTS

Papers containing inappropriately duplicated images

A total of 20,621 research papers containing the search term "western blot" from 40 different journals and 14 publishers were examined for inappropriate duplications of photographic images, with or without repositioning or evidence of alteration (Table S1). Of these, 8,138 (39.8%) were published by a single journal (*PLOS ONE*) in 2013 and 2014; the other 12,483 (60.5%) papers were published in 39 journals spanning the years 1995-2014 (Fig. 1). Overall, 782 (3.8%) of these papers were found to include at least one figure containing inappropriate duplications.

Classification of inappropriately duplicated images

Problematic images were classified into three major categories: simple duplications, duplications with repositioning, and duplications with alteration.

- Category I: Simple Duplications. Figures containing two or more identical panels, either within the same figure or between different figures within the same paper, purporting to represent different experimental conditions, were classified as simple duplications. The most common examples in this category were beta-actin loading controls that were used multiple times to represent different experiments or identical microscopy images purporting to be obtained from different experiments. For papers containing such figures, the methods and results were reviewed to establish that the duplicated figures were indeed re-used for different experiments. The re-use of loading controls in

114 different figures obtained from the same experiment was not considered to be a
115 problem. Examples of simple duplication are shown in Fig. 2.

116 - Category II: Duplication with Repositioning. This category included microscopic or blot
117 images with a clear region of overlap, where one image had been shifted, rotated, or
118 reversed in respect to the other. Fig. 3 shows examples of duplicated figures with
119 repositioning.

120 - Category III: Duplication with Alteration. This category consisted of images that were
121 altered with complete or partial duplication of lanes, bands, or groups of cells,
122 sometimes with rotation or reversal in respect to each other, within the same image
123 panel or between panels or figures. This category also includes figures containing
124 evidence of "stamping" in which a defined area is duplicated multiple times within the
125 same image, "patching" in which part of an image is obscured by a rectangular area of
126 different background, and FACS images sharing conserved regions and other regions in
127 which some data points have been added or removed. Examples of duplicated images
128 with alteration are shown in Fig. 4.

129

130 Two additional types of image modification were not scored as problematic, although they may
131 represent questionable research practices by current standards and would not be accepted by
132 certain journals (i.e., *Journal of Cell Biology*):

133 - Cuts. Abrupt vertical changes in the background signal between adjacent lanes in a blot
134 or gel suggest that the lanes were not next to each other in the original gel. These

splices are of potential concern but do not necessarily indicate inaccurate data representation.

- Beautification. Part of the background of a blot or gel where no band of interest is expected may show signs of patching, perhaps to remove a smudge or stain. This is not considered to represent best practice according to contemporary guidelines for data presentation (14) but does not necessarily indicate inaccurate data representation.

Although researcher intent could not be definitively determined in this study, the three categories of duplicated images were felt to have different implications with regard to the likelihood of scientific misconduct. Category I (simple duplication) images are most likely to result from honest errors, in which an author intended to insert two similar images but mistakenly inserted the same image twice. Alternatively, simple duplications may result from misconduct, for example, if an author intentionally recycled a control panel from a different experiment because the actual control was not performed. Category II (duplication with repositioning) and category III (duplication with alteration) may be somewhat more likely to result from misconduct, as conscious effort would be required for these actions.

In our study, a paper was classified as containing an inappropriate duplication when at least one category I, II, or III problem was identified. Papers were classified according to the highest category of duplicated image; e.g., a paper containing both category I and category II images was classified as a category II paper.

Among the 782 problematic papers found in this study, 230 (29.4%) contained simple duplications, 356 (45.5%) contained duplicated images with repositioning, while the remaining 196 (25.1%) contained duplicated figures with alteration.

157

158 **Temporal trends in image duplication**

159 To investigate the prevalence of image duplications and alterations over time, we plotted the
160 percentage of papers containing inappropriate image duplication as a function of publication
161 year (Fig. 5). The percentage of papers with image duplications appeared to be relatively low (<
162 2%) from 1995-2002, with no problematic images found among the 194 papers screened from
163 1995. However, a sharp increase in the percentage of papers with duplicated images was
164 observed in the year 2003 (3.6%), after which the percentages have remained close to or above
165 4%. This pattern remained very similar when only a subset of 16 journals for which papers were
166 scanned from all 20 years was considered, except for a decline in the duplications found in 2014,
167 the last year of our screen (2.2%) (Fig. 5).

168

169 **Correlation of impact factor with image duplication**

170 Substantial variation in the prevalence of papers with image duplication was observed among
171 the 40 journals investigated. In *PLOS ONE*, from which the largest number of papers was
172 screened, 4.3% of the papers were found to contain inappropriately duplicated images,
173 whereas the percentage of papers with image duplication ranged from 0.3% (*Journal of Cell*
174 *Biology)* to 12.4% (*International Journal of Oncology*) among the other journals, with a mean of
175 4.4% in the last decade. Hence, even though *PLOS ONE* was the journal that provided the
176 largest set of papers evaluated in this study, it is not an outlier with regard to inappropriately
177 duplicated images relative to the other journals examined. To assess the possibility that
178 journals with higher impact factors might be better at detecting problematic images and/or

179 authors might be more careful in preparing images for publication in such journals, the
180 relationship between the prevalence of image duplication and journal impact factor was
181 examined (Fig. 6). For this analysis, only papers published between 2005 and 2014 were
182 included, because the prevalence of problematic images was lower in older publications, and
183 older papers were only evaluated for selected journals. A negative correlation between image
184 duplication and journal impact factor was observed (Pearson's correlation; p-value 0.019), with
185 the lowest percentage of problematic images found in journals with high impact factors. The
186 prevalence of image duplication in 12 open access journals was not significantly different from
187 that in 28 non-open access journals (p = 0.38, chi-squared test).
188
189 **Country of origin of papers containing image duplication**
190 To determine whether inappropriate image duplication was more frequent in some countries
191 than in others, the country of origin for each of the 348 papers from *PLOS ONE* containing
192 duplicated images was compared to the country of origin for all papers published by that
193 journal during the same time interval that were included in our search. In cases where the
194 authors of a paper were affiliated with institutions in multiple countries, all countries were
195 taken into account. A majority of the 8,138 screened papers published in *PLOS ONE* during the
196 16-month study period from 2013-2014 were affiliated with China (26.2%) and the US (40.9%)
197 (Fig. 7). However, papers from China had a 1.89-fold higher probability of containing
198 problematic images than would have been predicted from the frequency of publication (chi-
199 squared test, p-value < 0.001), while papers from the US had a lower probability (0.60-fold)
200 (chi-squared test, p-value < 0.001). Other countries with a higher-than-predicted ratio of

papers containing image duplication were India (1.93) and Taiwan (1.20), whereas the prevalence of image duplication was lower than predicted in papers from the UK (0.47), Japan (0.26) and Germany (0.34).

Number of authors and image duplication

Errors or misconduct might be predicted to be more frequent in papers with fewer authors, due to reduced scrutiny. The mean number of authors per paper for the 781 papers containing inappropriate image duplication was 7.28. No significant difference between the mean number of authors for papers with or without image duplication was found ($p > 0.1$).

Problematic images in multiple papers by the same author

Our dataset of 782 problematic papers contained 28 papers (i.e., 14 pairs) of papers with a common first author. To determine whether authors of papers containing inappropriate image duplication were more likely to have published additional papers containing image duplication, we screened other papers written by the first and last authors of 559 papers (all from unique first authors) identified during initial screening. This analysis encompassed 2,425 papers, or a mean of 4.3 additional papers for each primary paper. In 217 cases (38.8%), at least one additional paper containing duplicated images was identified. In total, 269 additional papers containing duplicated images (11.1%) were found out of the 2,425 papers in the secondary dataset. The percentage of papers with duplicated images in the secondary dataset was significantly higher than that of the first dataset (11.1% vs. 3.8%, chi-squared test: $p<0.001$),

222 indicating that other papers by first or last authors of papers with duplicated images have an
223 increased probability of also containing duplicated images.
224
225 **DISCUSSION**
226
227 The quality and integrity of the scientific literature has been increasingly questioned (8,
228 9, 15, 16). The present study attempts to empirically determine the prevalence of one type of
229 problematic data, inappropriate image duplication, by visual inspection of electrophoretic,
230 microscopic and flow cytometric data from more than 20,000 recent papers in 40 primary
231 research journals. The major findings of this study are: (1) figures containing inappropriately
232 duplicated images can be readily identified in published papers through visual inspection
233 without the need for special forensic software methods or tools; (2) approximately 1 of every
234 25 published papers contains inappropriately duplicated images; (3) the prevalence of papers
235 with inappropriate image duplication rose sharply after 2002 and has since remained at
236 increased levels; (4) the prevalence of inappropriate image duplication varies among journals
237 and correlates inversely with journal impact factor; (5) papers containing inappropriately
238 duplicated images originated more frequently from China and India and less frequently from
239 the US, UK, Germany, Japan, or Australia; and (6) other papers by authors of papers containing
240 inappropriately duplicated images often contained duplicated images as well. These findings
241 have important implications for the biomedical research enterprise and suggest a need to
242 improve the literature though greater vigilance and education.

243 The finding that figures with inappropriate duplications can be readily identified by
244 simple inspection suggests that greater scrutiny of publications by authors, reviewers, and
245 editors might be able to identify problematic figures prior to publication. The *Journal of Cell*
246 *Biology* was among the first to call attention to the problem of figure alteration in manuscripts
247 (17), and this journal instituted a policy to carefully inspect all manuscripts for image
248 manipulation prior to publication (14). The low prevalence of problematic images in this
249 journal (0.3%) suggests that these measures have been effective. The *EMBO Journal*, which was
250 not part of the present study, has also instituted a manual screening process for aberrant
251 images (18). Our findings are consistent with the notion that greater scrutiny by journals can
252 reduce the prevalence of problematic images. However, this is likely to require a concerted
253 effort by all journals, so that authors of papers with problematic data do not simply avoid
254 publication in venues that employ rigorous screening procedures.

255 The prevalence of papers containing inappropriate image duplication increased
256 markedly in 2003 and has remained high in subsequent years. This coincides with the observed
257 increase in retracted publications (8) and provides empirical evidence that the increased
258 prevalence of problematic data is not simply a result of increased detection, as has been
259 suggested (19). Although the causes of the increased frequency of image duplication since
260 2003 are not known, we have considered several possible explanations. First, older papers
261 often contain figures with lower resolution, which may have obscured evidence of manipulation.
262 Second, the widespread availability and usage of digital image modification software in recent
263 years may have provided greater opportunity for both error and intentional manipulation.
264 Third, the increasing tendency for images to be directly prepared by authors instead of by

265 professional photographers working in consultation with authors has removed a potential
266 mechanism of quality control. One possible mechanism to reduce errors at the laboratory level
267 would be to involve multiple individuals in the preparation of figures for publication. The lack
268 of correlation between author number and the frequency of image duplication suggests that
269 the roles of most authors are compartmentalized or diluted, such that errors or misconduct are
270 not readily detected. A fourth consideration is that increasing competition and career-related
271 pressures may be encouraging careless or dishonest research practices (20). Finally, electronic
272 manuscript submission, as implemented by many journals in the early 2000s, facilitated
273 submissions from countries that were previously discouraged to submit because of high postal
274 costs.

275 A large variation in the prevalence of papers containing inappropriately duplicated
276 images was observed among journals ranging from the *Journal of Cell Biology* (0.3%) to the
277 *International Journal of Oncology* (12.4%), a difference of more than 40-fold. The differences
278 among journals are important because, as noted above, these suggest that journal editorial
279 policies can have a substantial impact on this problem. Alternatively, the variable prevalence of
280 duplication could be partly accounted for by variations in the average number of figures and
281 panels-per-figure, which is likely to differ per journal but was not determined in our study. The
282 inverse correlation between the prevalence of problematic papers and journal impact factor
283 contrasts with the positive correlation observed for research misconduct resulting in retraction
284 (8, 21–23). Although the association was weak, this may suggest that higher impact journals
285 are better able to detect anomalous images prior to publication. Alternatively, authors

286 submitting to such journals may be more careful with figure preparation. Nevertheless, we
287 note that even the most highly selective journals contain some papers with figures of concern.

288 China and the United States were responsible for the majority of papers containing
289 inappropriately duplicated images, which is not surprising given the large research output of
290 these countries. However it is noteworthy that the proportion of *PLOS ONE* papers from China
291 and India that were found to contain problematic images was higher than would be predicted
292 from their overall representation, whereas the opposite was true for papers from the US, UK,
293 Germany, Japan and Australia. This suggests that ongoing efforts at scientific ethics reform in
294 China and India should pay particular attention to issues relating to figure preparation (24, 25).
295 The analysis of geographic origin was limited to papers published in *PLOS ONE*, because this
296 journal offered an online tool to search for this information. The geographic distribution of
297 papers with problematic images may be different in other journals.

298 In nearly 40% of the instances in which a problematic paper was identified, screening of
299 other papers from the same authors revealed additional problematic papers in the literature.
300 This suggests that image duplication results from systematic problems in figure preparation by
301 individual researchers, which tend to recur.

302 Our findings suggest that as many as 1 out of every 25 published papers containing
303 Western blot or other photographic images could contain data anomalies. This is likely to be an
304 underestimate of the extent of problematic data in the literature for several reasons. First, only
305 image data were analyzed, thus errors or manipulation involving numerical data in graphs or
306 tables would not have been detected. Second, only duplicated images within the same paper
307 were examined, thus the reuse of images in other papers by the same author(s) would not have

308 been detected. Third, since problematic images were detected by visual inspection, the false-
309 negative rate could not be determined; we readily acknowledge that our screen may have
310 missed problematic papers. It should be noted that our findings contrast with a recent small
311 study by Oksvold that examined 120 papers from 3 different cancer research journals, reporting
312 duplicated images in 24.2% of the papers examined (26). Many of the reported image
313 duplications in the Oksvold study involved representation of identical experiments, which we
314 do not regard as necessarily inappropriate, as this form of duplication does not alter the
315 research results. For comparison, we screened 427 papers from the same 3 journals examined
316 by Oksvold and found the average percentage of problematic papers in these journals to be 6.8,
317 which is closer to our findings for other journals. Moreover, our study included more than
318 20,000 papers from 40 journals; in addition to more rigorous inclusion criteria, we required
319 consensus between three independent examiners for an image to be classified as containing
320 inappropriate duplication, ensuring a low false-positive rate.

321 The high prevalence of inaccurate data in the literature should be a finding of
322 tremendous concern to the scientific community, since the literature is the record of scientific
323 output upon which future research progress depends. Papers containing inaccurate data can
324 reduce the efficiency of the scientific enterprise by directing investigators to pursue false leads
325 or construct unsupportable hypotheses. Although our findings are disturbing, they also suggest
326 specific actions that can be taken to improve the literature. Increased awareness of recurring
327 problems with figure preparation, such as control band duplication, can lead to the reform of
328 laboratory procedures to detect and correct such issues prior to manuscript submission. The
329 variation among journals in the prevalence of problematic papers suggests that individual

330 journal practices, such as image screening, can reduce the prevalence of problematic images
331 (17, 18). The problems identified in this study provide further evidence for the scientific
332 establishment that current standards are insufficient to prevent flawed papers from being
333 published. Our findings call for the need of greater efforts to ensure the reliability and
334 integrity of the research literature.
335
336 **MATERIALS AND METHODS**
337
338 **Selection strategy**
339 A total of 20,621 papers were selected from 40 different scientific journals in the fields of
340 microbiology and immunology, cancer biology, and general biology. These journals were
341 published by 14 organizations (average of 2.9 journals per publisher, range 1-6) (Table S1). All
342 journals included in the search were indexed in PubMed, with a mean impact factor of 6.9
343 (range 1.3 - 42.4, Thomson Reuters 2013). Papers were examined if they contained the search
344 term “western blot," using the search tool provided at the journal's website. Only original
345 research papers containing figures were included; retracted papers, review papers, and
346 conference abstracts were excluded. Corrected papers were included only if the correction
347 involved issues other than the problems identified by the present analysis (e.g., incorrect grant
348 statement). From a single journal (*PLOS ONE*), 8,138 papers published in 2013 and 2014 were
349 included in the study, comprising 39.5% of the dataset. From the remaining journals, a mean of
350 320.1 (range 77 - 1070) papers per journal were included. For most of these journals, if more
351 than 50 papers were found in a given year, screening was limited to the first 40-50 papers that

352 were shown in the search field. The selected papers spanned the years 1995-2014, with most
353 papers published in 2013-2014, primarily as a result of the large contribution of papers from
354 *PLOS ONE* (Fig. 1). The large number of *PLOS ONE* papers analyzed reflects both the journal
355 format, which facilitates image analysis, and the fact that *PLOS ONE* is currently the world's
356 largest scientific journal with approximately 30,000 new articles per year
357 (https://en.wikipedia.org/wiki/PLOS_ONE).
358

359 **Visual screening**

360 All papers were screened by examining images at the publisher's website. Although papers
361 were selected using the search term “western blot," all types of photographic image were
362 examined, including protein and nucleic acid electrophoretic gels and blots, histology or
363 microscopy images, and photos of culture plates, plants, and animals. Fluorescence-activated
364 cell sorting (FACS) plots were included as well since these, like photographic images,
365 purportedly represent raw data. Figure panels containing line art such as bar graphs or line
366 graphs were not included in the study. Images within the same paper were visually inspected
367 for inappropriate duplications, repositioning, or possible manipulation (e.g., duplications of
368 bands within the same blot). All papers were initially screened by one of the authors (EMB). If
369 a possible problematic image or set of images was detected, figures were further examined for
370 evidence of image duplication or manipulation using the Adjust Color tool in Preview software
371 on an Apple iMac computer. No additional special imaging software was used. Supplementary
372 figures were not part of the initial search but were examined in papers in which problems were
373 found in images in the primary manuscript. All figures found by the screening author (EMB) to

374 contain possible duplications were independently reviewed by the two co-authors (FCF and AC).
375 Consensus had to be reached among all three authors for a paper to be considered to contain
376 unequivocal evidence of inappropriate figure duplication. Consensus among all three authors
377 was reached in 90.4% of the papers selected during primary screening.
378
379 **Statistical analysis**
380 The R software package was used to plot and analyze data. The stat_smooth method
381 implemented in the R ggplot2 library was applied to find the best fit for a linear regression
382 model on semi-log transformed data examining the relationship between the 2013 Thomson
383 Reuters impact factor and percentage of papers with inappropriately duplicated images for the
384 40 journals included in this study. Pearson's correlation and Pearson's Chi-square tests with
385 Yates' continuity correction were performed in basic R. R code is available as Supplemental
386 Data S1.
387
388 **FUNDING INFORMATION**
389
390 This research received no specific grant from any funding agency in the public, commercial, or
391 not-for-profit sectors.
392
393 **SUPPLEMENTAL MATERIAL**
394 1. Supplemental Table S1: The 40 journals screened in this study
395 2. Supplemental Data S1: R code

396 **REFERENCES**

397

398 1. **Casadevall A, Steen RG, Fang FC**. 2014. Sources of error in the retracted scientific
399 literature. FASEB J **28**:3847–3855.

400 2. **Steneck NH**. 2006. Fostering integrity in research: definitions, current knowledge, and
401 future directions. Sci Eng Ethics **12**:53–74.

402 3. **Fanelli D**. 2009. How many scientists fabricate and falsify research? A systematic review
403 and meta-analysis of survey data. PLoS ONE **4**:e5738.

404 4. **Lüscher TF**. 2013. The codex of science: honesty, precision, and truth--and its violations.
405 Eur Heart J **34**:1018–1023.

406 5. **Geison GL**. 1996. The Private Science of Louis Pasteurillustrated, revised. Princeton
407 University Press.

408 6. **Alberts B, Cicerone RJ, Fienberg SE, Kamb A, McNutt M, Nerem RM, Schekman R,**
409 **Shiffrin R, Stodden V, Suresh S, Zuber MT, Pope BK, Jamieson KH**. 2015. SCIENTIFIC
410 INTEGRITY. Self-correction in science at work. Science **348**:1420–1422.

411 7. **Steen RG, Casadevall A, Fang FC**. 2013. Why has the number of scientific retractions
412 increased? PLoS ONE **8**:e68397.

413 8. **Fang FC, Steen RG, Casadevall A**. 2012. Misconduct accounts for the majority of retracted
414 scientific publications. Proc Natl Acad Sci U S A **109**:17028–17033.

415 9. **Begley CG, Ellis LM**. 2012. Drug development: Raise standards for preclinical cancer
416 research. Nature **483**:531–533.

417 10. **Prinz F, Schlange T, Asadullah K**. 2011. Believe it or not: how much can we rely on

418 published data on potential drug targets? Nat Rev Drug Discov **10**:712.

419 11. **Open Science Collaboration**. 2015. PSYCHOLOGY. Estimating the reproducibility of
420 psychological science. Science **349**:aac4716.

421 12. **Flaherty DK**. 2011. The vaccine-autism connection: a public health crisis caused by
422 unethical medical practices and fraudulent science. Ann Pharmacother **45**:1302–1304.

423 13. **Bowen A, Casadevall A**. 2015. Increasing disparities between resource inputs and
424 outcomes, as measured by certain health deliverables, in biomedical research. Proc Natl
425 Acad Sci U S A.

426 14. **Rossner M, Yamada KM**. 2004. What's in a picture? The temptation of image
427 manipulation. J Cell Biol **166**:11–15.

428 15. **Ioannidis JP**. 2005. Why most published research findings are false. PLoS Med **2**:e124.

429 16. **Collins FS, Tabak LA**. 2014. Policy: NIH plans to enhance reproducibility. Nature **505**:612–
430 613.

431 17. **Rossner M**. 2002. Figure manipulation: assessing what is acceptable. J Cell Biol **158**:1151–
432 1151.

433 18. **Pulverer B**. 2014. EMBO Press - a new way to publish. EMBO J **33**:1–2.

434 19. **Fanelli D, Costas R, Larivière V**. 2015. Misconduct Policies, Academic Culture and Career
435 Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. PLoS ONE
436 **10**:e0127556.

437 20. **Anderson MS, Ronning EA, De Vries R, Martinson BC**. 2007. The perverse effects of
438 competition on scientists' work and relationships. Sci Eng Ethics **13**:437–461.

439 21. **Fang FC, Casadevall A**. 2011. Retracted science and the retraction index. Infect Immun

440 **79**:3855–3859.

441 22. **Cokol M, Iossifov I, Rodriguez-Esteban R, Rzhetsky A**. 2007. How many scientific papers
442 should be retracted? EMBO Rep **8**:422–423.

443 23. **Liu SV**. 2006. im1.biz/TopRetraction.pdf. Scientific Ethics **1**:91–93.

444 24. **Yidong G, Xin H**. 2006. Research ethics. China's science ministry fires a barrage of
445 measures at misconduct. Science **312**:1728–1729.

446 25. **Patnaik PR**. 2015. Scientific Misconduct in India: Causes and Perpetuation. Sci Eng Ethics.

447 26. **Oksvold MP**. 2015. Incidence of Data Duplications in a Randomly Selected Pool of Life
448 Science Publications. Sci Eng Ethics.

449 27. **Zou X, Sorenson BS, Ross KF, Herzberg MC**. 2013. Augmentation of epithelial resistance
450 to invading bacteria by using mRNA transfections. Infect Immun **81**:3975–3983.

451 28. **Zou X, Sorenson BS, Ross KF, Herzberg MC**. 2015. Correction for zou et Al., augmentation
452 of epithelial resistance to invading bacteria by using mRNA transfections. Infect Immun
453 **83**:1226–1227.

454 29. **Liu J, Xu P, Collins C, Liu H, Zhang J, Keblesh JP, Xiong H**. 2013. HIV-1 Tat protein
455 increases microglial outward K(+) current and resultant neurotoxic activity. PLoS ONE
456 **8**:e64904.

457 30. **The PLOS ONE Staff**. 2014. Correction: HIV-1 Tat Protein Increases Microglial Outward K+
458 Current and Resultant Neurotoxic Activity. PLoS ONE **9**:e109218.

459 31. **Liu Y, Liu Y, Sun C, Gan L, Zhang L, Mao A, Du Y, Zhou R, Zhang H**. 2014. Carbon ion
460 radiation inhibits glioma and endothelial cell migration induced by secreted VEGF. PLoS
461 ONE **9**:e98448.

462 32. **Liu Y, Liu Y, Sun C, Gan L, Zhang L, Mao A, Du Y, Zhou R, Zhang H**. 2015. Correction:
463 Carbon Ion Radiation Inhibits Glioma and Endothelial Cell Migration Induced by Secreted
464 VEGF. PLoS ONE **10**:e0135508.
465 33. **Pulloor NK, Nair S, McCaffrey K, Kostic AD, Bist P, Weaver JD, Riley AM, Tyagi R, Uchil**
466 **PD, York JD, Snyder SH, García-Sastre A, Potter BV, Lin R, Shears SB, Xavier RJ, Krishnan**
467 **MN**. 2014. Human genome-wide RNAi screen identifies an essential role for inositol
468 pyrophosphates in Type-I interferon response. PLoS Pathog **10**:e1003981.
469 34. **The PLOS Pathogens Staff**. 2014. Correction: Human Genome-Wide RNAi Screen
470 Identifies an Essential Role for Inositol Pyrophosphates in Type-I Interferon Response.
471 PLoS Pathog **10**:e1004519.
472 35. **Xu X, Liu T, Zhang A, Huo X, Luo Q, Chen Z, Yu L, Li Q, Liu L, Lun ZR, Shen J**. 2012. Reactive
473 oxygen species-triggered trophoblast apoptosis is initiated by endoplasmic reticulum
474 stress via activation of caspase-12, CHOP, and the JNK pathway in Toxoplasma gondii
475 infection in mice. Infect Immun **80**:2121–2132.
476 36. 2015. Retraction for Xu et al., Reactive oxygen species-triggered trophoblast apoptosis is
477 initiated by endoplasmic reticulum stress via activation of caspase-12, CHOP, and the JNK
478 pathway in Toxoplasma gondii infection in mice. Infect Immun **83**:1735.
479 37. **Friedman-Levi Y, Mizrahi M, Frid K, Binyamin O, Gabizon R**. 2013. PrP(ST), a soluble,
480 protease resistant and truncated PrP form features in the pathogenesis of a genetic prion
481 disease. PLoS ONE **8**:e69583.
482 38. **Friedman-Levi Y, Mizrahi M, Frid K, Binyamin O, Gabizon R**. 2015. Correction: PrPST, a
483 Soluble, Protease Resistant and Truncated PrP Form Features in the Pathogenesis of a

484 Genetic Prion Disease. PLoS ONE **10**:e0133911.

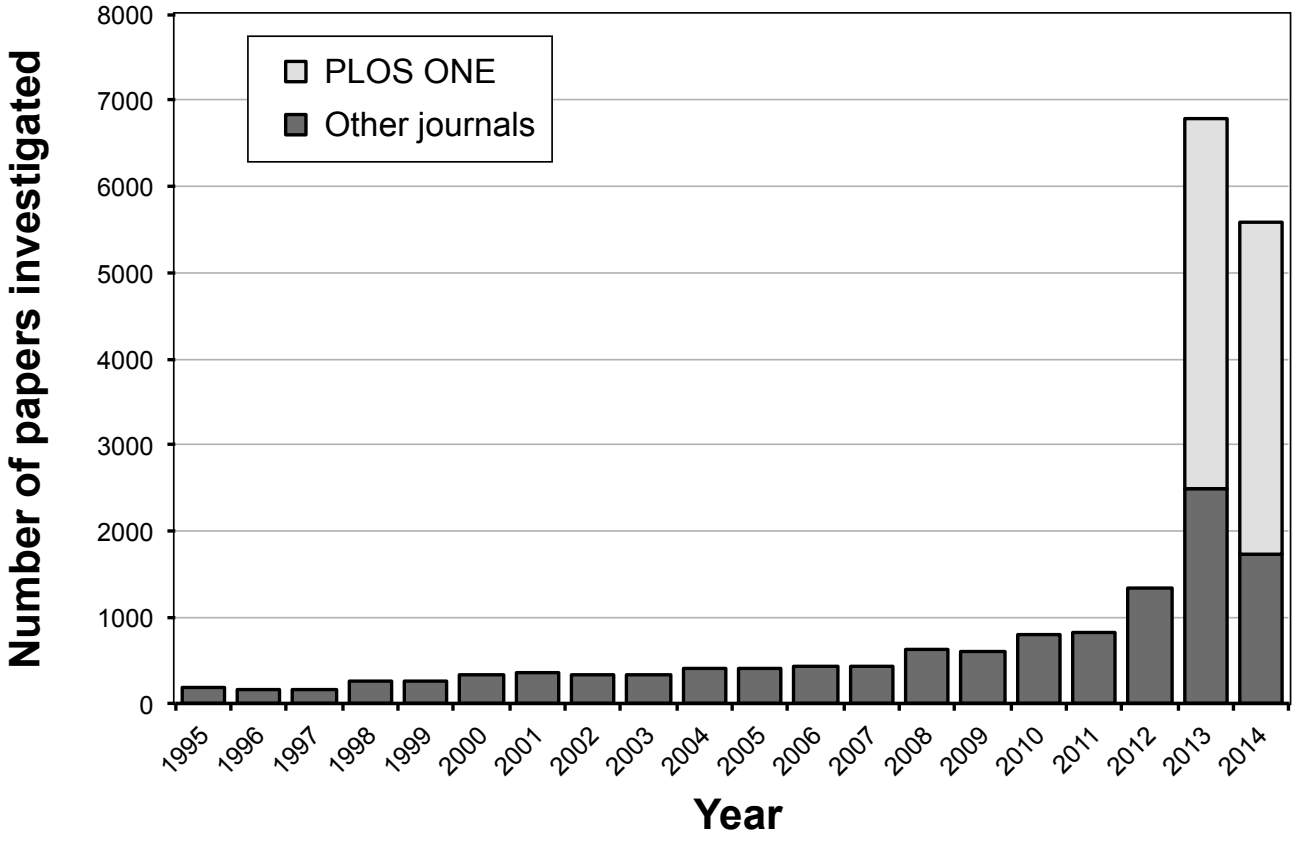
485

486

487

488

489 **FIGURES and LEGENDS**



490

491 **Figure 1. Publications investigated in this study by year of publication.** The majority of
492 screened papers were published in 2013 and 2014 due to the large proportion of *PLOS ONE*
493 papers (39.5%) in the dataset. The lowest number of papers (n=151) screened in this study was
494 published in 1996.

495

496

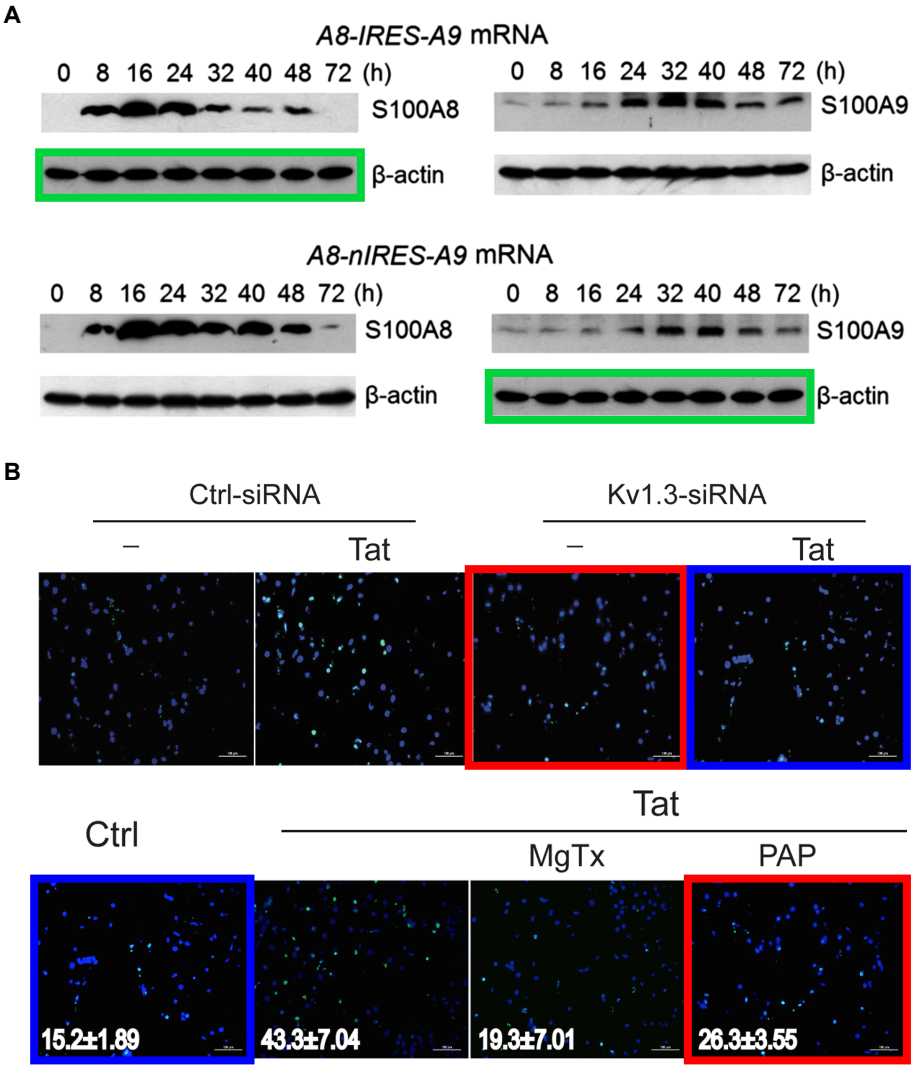
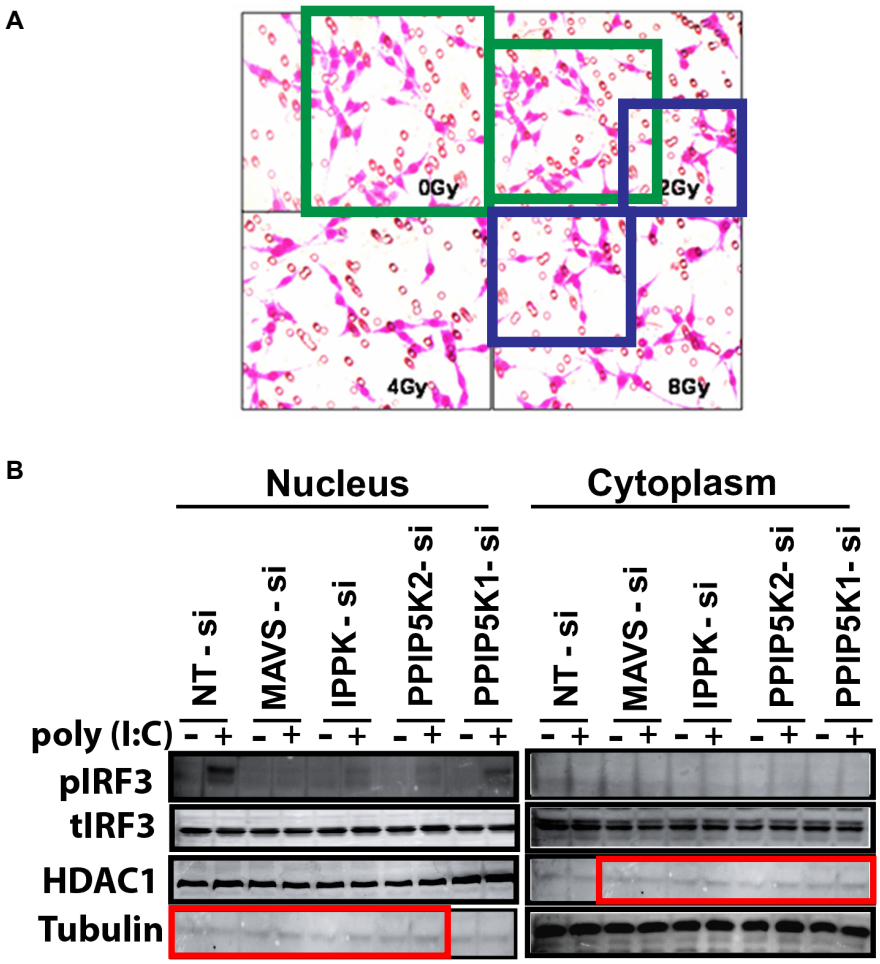


Figure 2. Examples of simple duplication (Category I). A. The beta-actin control panel in the top left is identical to the panel in the bottom right (shown by green boxes), although each panel represents a different experimental condition (27); corrected in: (28). Figure reproduced with permission from the publisher. **B.** The panels shown here were derived from two different figures within the same paper. Two of the top panels appear identical to two of the bottom panels, but they represent different experimental condition (shown with red and blue boxes) (29); corrected in: (30). Figure reproduced under the Creative Commons (CC BY) license. All duplications might have been caused by honest errors during assembly of the figures.



506

507 **Figure 3. Examples of duplication with repositioning (Category II). A.** Although the panels

508 represent four different experimental conditions, three of the four panels appear to show a

509 region of overlap (shown with green and blue boxes), suggesting that these photographs were

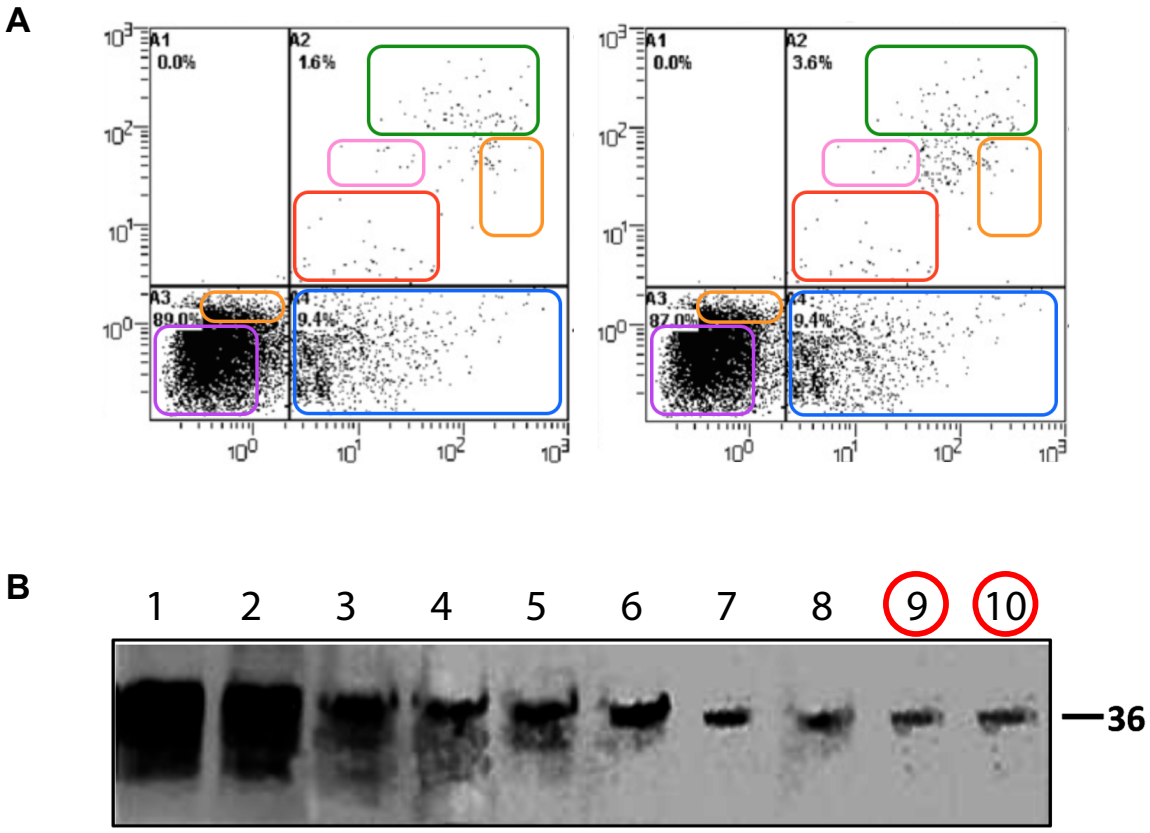
510 actually obtained from the same specimen (31); corrected in: (32)**. B.** Western blot panels

511 "Nucleus-Protein D" and "Cytoplasm Protein C" purportedly depict different proteins and

512 cellular fractions, but the blots appear very similar albeit shifted by two lanes (shown with red

513 boxes) (33); corrected in: (34). Both figures reproduced under the Creative Commons (CC BY)

514 license.



515

516

517 **Figure 4. Examples of duplication with alteration (Category III). A.** The left and right FACS

518 panels represent different experimental conditions and show different percentages of cell

519 subsets, but regions of identity (colored boxes) between the panels suggest that the images

520 have been altered (35); retracted in (36). Figure reproduced with permission from the publisher.

521 **B.** The figure shown here displays a Western blot of 10 different protein fractions isolated from

522 a density gradient. The figure appears to show a single blot, but the last two lanes (highlighted

523 with red circles) appear to contain an identical band. Exposure was altered to bring out details

524 (37); corrected in: (38). Figure reproduced under the Creative Commons (CC BY) license.

525

526

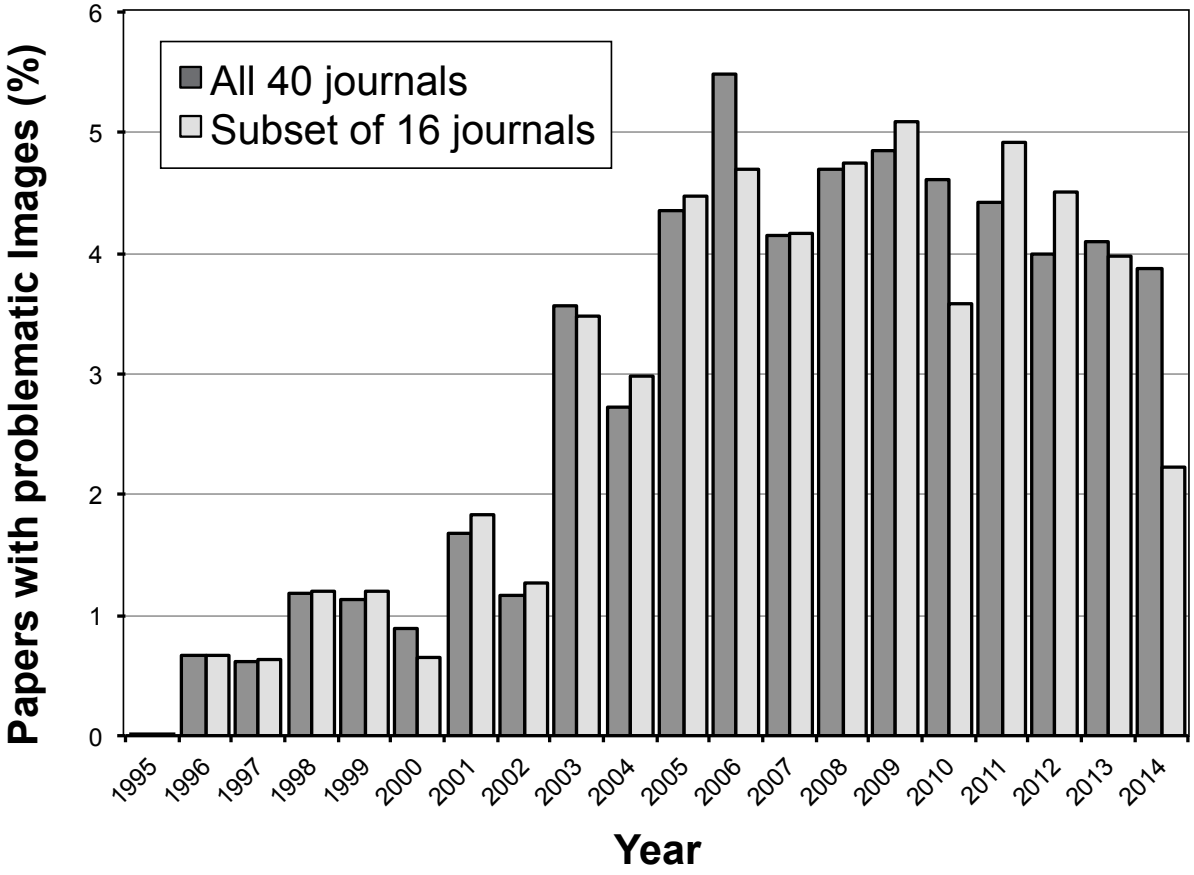
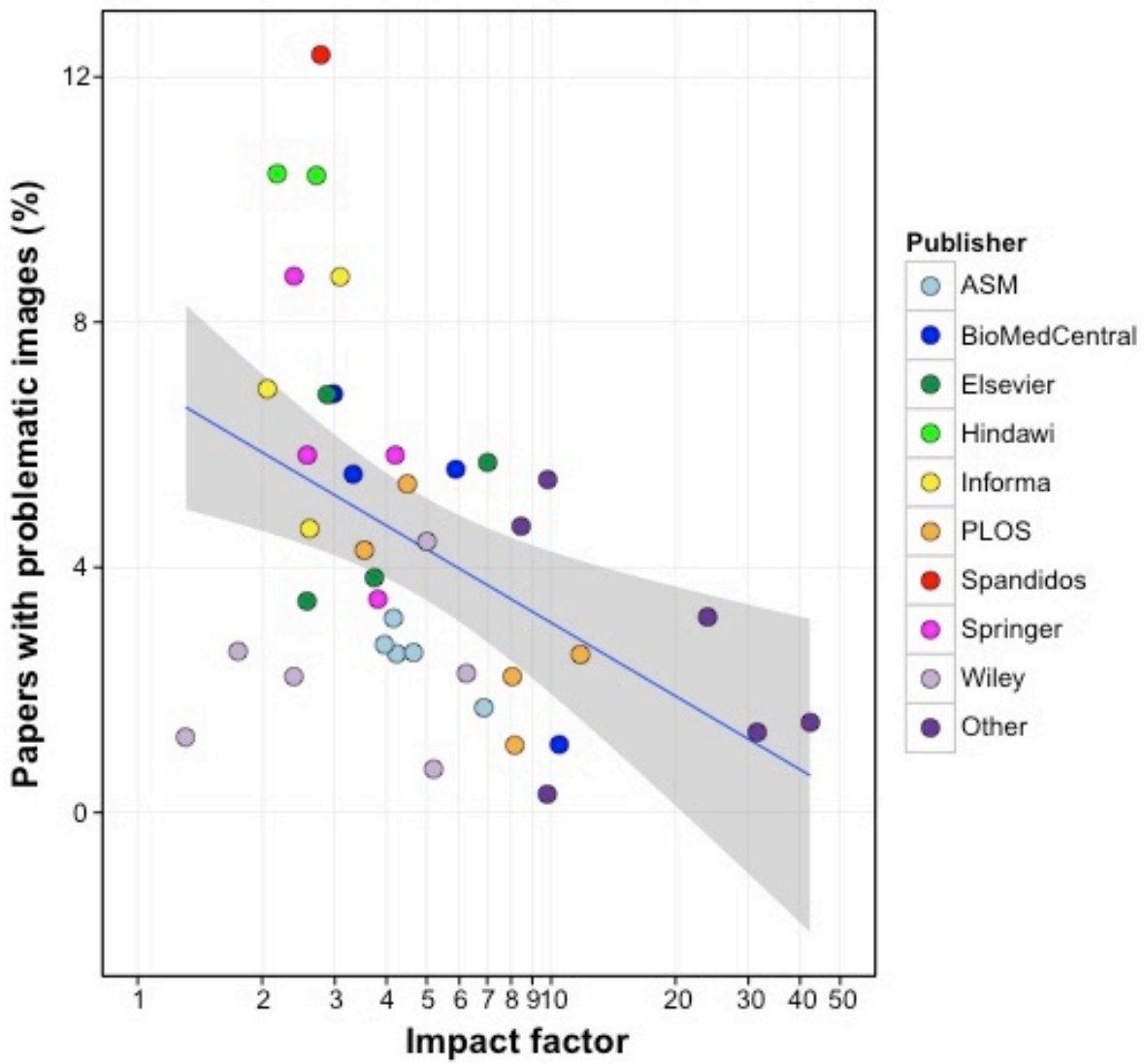


Figure 5. Percentage of papers containing inappropriate image duplications by year of publication. No papers with duplications were found in 1995. The dark gray bars show the data for all 40 journals. The light gray colored bars show a subset of 16 journals for which papers spanning the complete timespan of 20 years were scanned. The total numbers of papers screened in each year are shown in Figure 1.



535

536 **Figure 6. Correlation between journal impact factor and percentage of papers with image**
537 **duplication.** Only papers from 2005-2014 (n = 17,816) were included in this analysis. Each data
538 point represents a journal included in this study (n=40), with data points color-coded according
539 to their publisher (n=14; journals published by AAAS, Nature, Cell Press, the National Academy
540 of Sciences, and the Rockefeller University Press are grouped under "Other"). The x-axis is
541 shown on a logarithmic scale due to the small number of journals with a high impact factor
542 included in this study. The blue line shows a linear regression model. The grey zone depicts the
543 95% confidence interval.

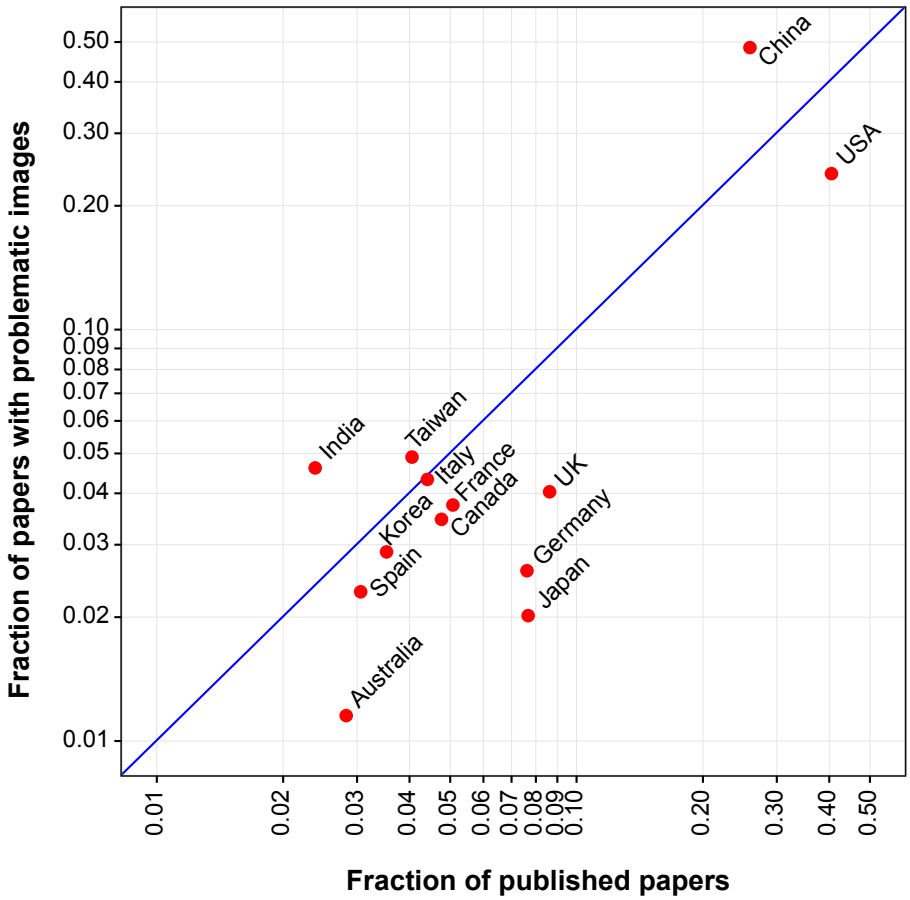


Figure 7. Proportion of papers with image duplications by country. The proportion of papers affiliated with specific countries submitted to *PLOS ONE* during a 16-month period in the years 2013 and 2014 (n = 8,138) is plotted versus the proportion of *PLOS ONE* papers from that same period containing inappropriate image duplication affiliated with specific countries (n = 348). Each data point represents a country for which 100 or more papers were screened. Some papers were affiliated with more than one country. The blue line represents the line where data points are expected to fall if problematic papers are distributed as expected according to their representation in the journal. Countries plotted above the blue line had a higher-than-expected proportion of problematic papers; countries plotted below the line had a lower-than-expected ratio.

SUPPLEMENTARY TABLE

Table S1. The 40 journals screened in this study. Table includes publisher, impact factor (Thomson Reuters 2013), number of papers containing the term "western blot" (WB) screened per year, and number of papers with inappropriate image duplication (IDs) found in that year.

Bik et al. Supplementary Table 1

Journal Title	Publisher	Impact Factor 2013	Screened	Papers with ID	% ID	2014		2013		2012		2011		2010		2009		2008		2007		2006		2005		2004		2003		2002		2001		2000		1999		1998		1997		1996		1995	
						WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID	WB	ID
PLOS ONE	PLOS	3.53	**8138**	**348**	**4.28**	3846	167	4292	181	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS Pathogens	PLOS	8.06	**406**	**9**	**2.22**	200	5	206	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS Genetics	PLOS	8.17	**362**	**4**	**1.10**	164	2	198	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS Biology	PLOS	11.77	**233**	**6**	**2.58**	36	2	42	0	47	1	47	1	61	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PLOS NTD	PLOS	4.49	**317**	**17**	**5.36**	72	3	79	4	53	3	37	3	42	3	17	0	15	1	2	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Journal of Clinical Microbiology	ASM	4.23	**595**	**11**	**1.85**	13	0	10	0	12	0	15	0	19	0	18	1	20	0	11	**0**	31	2	44	2	55	3	45	2	55	0	57	0	49	0	58	0	46	1	13	0	8	0	16	0
Applied and Environmental Microbiology	ASM	3.95	**292**	**8**	**2.74**	56	0	46	1	65	0	60	3	65	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
mBio	ASM	6.88	**175**	**3**	**1.71**	64	0	53	1	30	1	22	1	6	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infection and Immunity	ASM	4.16	**1070**	**30**	**2.80**	120	2	95	5	121	6	40	2	51	2	51	2	51	0	53	0	52	1	58	2	52	1	42	3	47	0	45	1	49	1	51	1	47	0	11	0	14	1	19	0
Journal of Virology	ASM	4.65	**421**	**11**	**2.61**	NA	NA	421	11	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
International journal of cancer	Wiley	5.01	**226**	**10**	**4.42**	40	0	186	10	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Clinical Microbiology and Infection	Wiley	5.20	**199**	**1**	**0.50**	9	0	6	0	10	1	11	0	14	0	23	0	17	0	14	0	21	0	15	0	15	0	7	0	5	0	11	0	6	0	3	0	5	0	4	0	1	0	2	0
Journal of Applied Microbiology	Wiley	2.39	**200**	**3**	**1.50**	15	0	19	0	8	0	14	0	11	1	12	0	19	0	23	1	6	1	8	0	9	0	7	0	11	0	14	0	6	0	9	0	5	0	4	0	NA	NA	NA	NA
Environmental Microbiology	Wiley	6.24	**189**	**5**	**2.65**	49	1	24	0	19	2	18	0	22	0	17	1	16	0	5	0	3	0	3	0	2	0	2	0	4	0	0	0	3	1	2	0	NA	NA	NA	NA	NA	NA	NA	NA
Microbiology and Immunology	Wiley	1.31	**358**	**3**	**0.84**	19	0	29	0	23	0	22	0	19	1	17	0	19	0	6	0	3	0	5	1	7	1	7	0	18	0	18	0	35	0	23	0	27	0	27	0	18	0	16	0
Letters in Applied Microbiology	Wiley	1.75	**123**	**2**	**1.63**	10	0	6	0	7	2	10	0	17	0	7	0	6	0	3	0	8	0	2	0	6	0	7	0	2	0	6	0	7	0	8	0	1	0	2	0	5	0	3	0
BioMed Research International	Hindawi	2.71	**77**	**8**	**10.39**	42	4	35	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Evid Based Compl Alternat Med	Hindawi	2.18	**96**	**10**	**10.42**	20	2	34	4	25	3	12	1	3	0	1	0	0	0	0	0	1	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BMC Microbiology	BMC	2.98	**340**	**23**	**6.76**	44	3	44	3	28	0	50	3	56	6	39	3	37	2	11	**0**	8	2	5	0	5	0	2	1	5	0	6	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Genome Biology	BMC	10.47	**105**	**1**	**0.95**	24	0	15	1	8	0	7	0	4	0	6	0	6	0	9	0	6	0	5	0	5	0	5	0	2	0	3	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Breast Cancer Research	BMC	5.88	**403**	**20**	**4.96**	58	4	38	2	49	5	49	0	36	2	24	0	31	1	25	2	12	2	35	2	13	0	4	0	3	0	5	0	17	0	4	0	NA	NA	NA	NA	NA	NA	NA	NA
BMC Cancer	BMC	3.32	**145**	**8**	**5.52**	NA	NA	145	8	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Diagn Microbiol Infect Dis	Elsevier	2.57	**115**	**3**	**2.61**	6	0	7	0	14	1	15	2	9	0	6	0	8	0	14	0	5	0	3	0	3	0	6	0	4	0	5	0	1	0	4	0	3	0	0	0	1	0	1	0
Lung Cancer	Elsevier	3.74	**317**	**11**	**3.47**	12	0	24	0	38	1	28	1	25	2	28	1	23	1	23	1	20	2	14	0	15	0	29	1	14	0	13	1	1	0	2	0	4	0	0	0	2	0	2	0
Cytokine	Elsevier	2.87	**464**	**28**	**6.03**	31	1	55	2	66	3	60	5	30	1	28	4	39	4	15	**1**	35	4	37	2	23	1	15	0	6	0	2	0	6	0	3	0	3	0	4	0	4	0	2	0
Journal of Autoimmunity	Elsevier	7.02	**150**	**6**	**4.00**	8	1	13	0	13	0	11	0	11	0	6	1	8	0	6	1	16	1	13	2	14	0	8	0	1	0	2	0	2	0	3	0	3	0	3	0	2	0	7	0
Appl Microbiol Biotechnol	Springer	3.81	**230**	**8**	**3.48**	78	3	91	5	61	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Breast Cancer Res Treatment	Springer	4.20	**206**	**12**	**5.83**	61	5	45	3	100	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cancer Chemotherapy and Pharmacolog	Springer	2.57	**542**	**29**	**5.35**	29	0	42	2	39	1	54	5	43	1	52	4	59	4	25	3	38	1	31	3	30	3	21	2	18	0	12	0	12	0	11	0	8	0	7	0	7	0	4	0
Molecular and Cellular Biochemistry	Springer	2.39	**800**	**43**	**5.38**	40	5	40	4	40	6	40	3	40	3	40	2	40	3	40	3	40	4	40	2	40	0	40	1	40	3	40	2	40	1	40	0	40	1	40	0	40	0	40	0
Growth Factors	Informa	3.09	**166**	**10**	**6.02**	0	0	4	0	11	0	12	1	12	1	19	1	19	2	13	2	3	1	10	1	5	1	4	0	3	0	7	0	7	0	5	0	13	0	10	0	4	0	5	0
Cancer Investigation	Informa	2.06	**220**	**13**	**5.91**	12	0	11	1	17	1	23	2	35	2	37	2	26	3	13	1	9	0	5	1	9	0	6	0	3	0	4	0	3	0	2	0	1	0	0	0	1	0	3	0
Leukemia & Lymphoma	Informa	2.61	**404**	**13**	**3.22**	54	1	26	2	39	1	26	0	29	3	17	1	20	1	23	2	16	1	9	0	12	0	12	1	20	0	23	0	15	0	10	0	20	0	9	0	11	0	13	0
International Journal of Oncology	Spandidos	2.77	**89**	**11**	**12.36**	NA	NA	51	6	38	5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Science	AAAS	31.48	**681**	**9**	**1.32**	37	0	35	1	32	0	31	0	41	0	34	1	38	1	51	1	36	1	47	0	58	1	43	1	56	1	59	1	52	0	4	0	3	0	5	0	8	0	11	0
Nature	Nature	42.35	**750**	**12**	**1.60**	50	0	50	1	50	1	50	1	50	0	50	2	50	2	50	0	50	0	25	0	25	0	25	0	25	0	25	1	25	0	25	2	25	1	25	1	25	0	50	0
Nature Oncogene	Nature	8.46	**150**	**7**	**4.67**	50	2	50	4	50	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cancer Cell	Cell Press	23.89	**188**	**6**	**3.19**	55	2	63	2	70	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Journal of Cell Biology	RU Press	9.79	**329**	**1**	**0.30**	102	0	110	0	117	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PNAS	NAS	9.81	**350**	**19**	**5.43**	50	1	50	4	50	2	50	2	50	3	50	3	50	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Sum		**20621**	**782**		**5576**	**216**	**6790**	**278**	**1350**	**54**	**814**	**36**	**801**	**37**	**599**	**29**	**617**	**29**	**435**	**18**	**419**	**23**	**414**	**18**	**403**	**11**	**337**	**12**	**342**	**4**	**357**	**6**	**336**	**3**	**267**	**3**	**254**	**3**	**164**	**1**	**151**	**1**	**194**	**0**
	% ID				**3.79**	**3.87**		**4.09**		**4.00**		**4.45**		**4.52**		**4.74**		**4.41**		**4.14**		**5.49**		**4.35**		**2.73**		**3.56**		**1.17**		**1.68**		**0.89**		**1.12**		**1.18**		**0.61**		**0.66**		**0.00**	