

One in 25 papers contains inappropriately duplicated images, screen finds

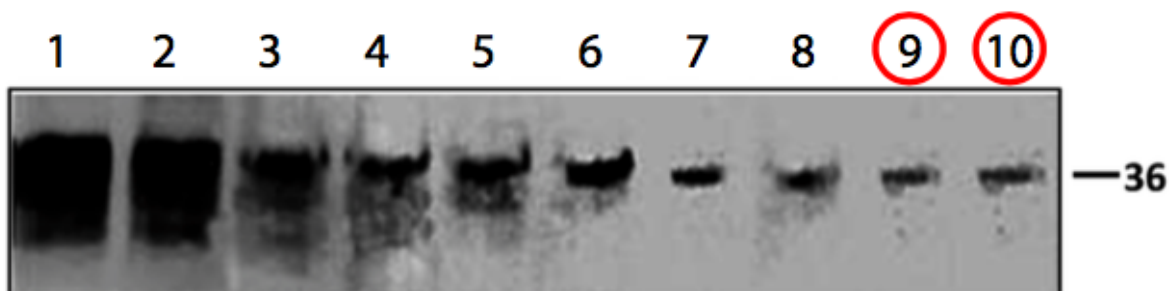
Elisabeth Bik, a microbiologist at Stanford, has for years been a behind-the-scenes force in scientific integrity, anonymously submitting reports on plagiarism and image duplication to journal editors. Now, she's ready to come out of the shadows.



Elies Bik

With the help of two editors at microbiology journals, she has conducted a massive study looking for image duplication and manipulation in 20,621 published papers. Bik and co-authors Arturo Casadevall and Ferric Fang (a board member of our parent organization) found 782 instances of inappropriate image duplication, including 196 published papers containing “duplicated figures with alteration.” The study is being released as a pre-print on bioArxiv.

An example the paper uses of “duplication with alteration” is this Western blot where a band has been duplicated:



Bik's procedure to find these kinds of duplications is disarmingly simple. She pulls up all the figures in a paper and scans them. It only takes her about a minute to check all the images in a *PLoS ONE* paper, a little longer for a paper with more complicated figures. In some cases, Bik adjusted the contrast on the image to better spot manipulations.

My initial screen is without software, just quickly flipping through images. When I see something that might be a duplication, I will download the figure(s) and inspect them using Preview (Mac software). I use simple color or brightness adjustments in Preview. I don't use Photoshop or other advanced programs.

It gets easier to spot problems the more you look, Bik told us on the phone:

In Western blots, every band has their own characteristics, they're like faces. I think if you train people, immediately they will see something is wrong. It takes you less than a second to recognize a face.

After screening the papers for this study, she then puts together detailed reports on the duplications and sends them to Casadevall and Fang, who both have to agree there was inappropriate duplication before inclusion in the paper.

For me, it's very obvious. It sort of shouts out to me. Some of these examples are almost funny, in a disturbing way – they're so obviously copied...On the other hand, these papers have been reviewed and seen by editors, and downloaded hundreds or thousands of times.

She admitted that there are almost certainly examples that she missed, noting that she got better and faster over time. But she said false positives are less likely, because the team required consensus from all three members to include a paper.

About 10% of the papers I flagged, we didn't agree so I took them out... When three people agree that an image looks very similar, it might still be a different image, but I think it's reason to flag that something is possibly wrong, and then talk to the author.

In 2013, *Nature* reported the findings of a previous screen of published papers — done by Italian bioinformatics startup BioDigitalValley founded by Enrico Bucci, and focusing mostly on gel-electrophoresis in Italian studies — which found about one in four papers had inappropriate duplications of images. Those included repeated use of the same image, as well as copy and pasted gel bands.

Mike Rossner of Image Data Integrity, former managing editor of the *Journal of Cell Biology*, was impressed by how many papers Bik screened, and how fast she was able to spot problems. “To look [for duplications] between different figures takes really good visual memory, and she must have that,” he told Retraction Watch. (Note: Rossner spoke to us about his new company in February.)

Rossner added that he wished the paper had compared its results to those available from journals that screen figures in accepted papers for manipulation, such as the *JCB* and *EMBO Journal*.

None of the papers included in Bik’s analysis had been retracted at the time she screened them. She has since submitted over 700 reports to journal editors showing the duplications, and written to about 10 institutions where she found what she calls “clusters” — three to six papers from the same group containing duplications.

This has resulted in six retractions, four of them in Fang’s journal, *Infection and Immunity* (which we covered at at the time) — much lower than the 42% retraction rate she has received in the past from reporting plagiarism she found through Google Scholar searches. About 60 inappropriately duplicated images have been corrected since Bik first began reporting these findings to editors in the spring of 2014. She estimates an average of six months between her reporting and the retractions or corrections over the course of this project.

I do plan to write another paper on the lack of outcome with these reports...It’s not in [journals’] interest to retract the papers, because it will bring down their impact factor. If we publish their response rate, then maybe we can motivate them to show they care about the science.

Bik has struggled with journals in another respect — getting this latest analysis published. The paper has been rejected by three journals, one after peer review. The other two journals chose not to send it out at all.

“I expect this to be a controversial paper, no journal wants to hear a percentage of their papers is considered very bad,” Bik told Retraction Watch when asked why she thinks it was rejected. “One reviewer said, oh, this has to be published. Most of the others said it’s very controversial, that it’s not novel.”

At one of the journals that didn’t send it for peer review, the editors told Bik it was important work, but not suitable for their audience.

“The fact that image manipulation is going on is a problem with the scientific record,” Rossner told Retraction Watch. “It seems to me she should be able to find a place that will publish this.”

Casadevall is the editor in chief of *mBio*, which showed an inappropriate duplication rate of 1.71%; at *Infection and Immunity*, that rate was 2.8%.

David Vaux, a cell biologist at the Walter + Eliza Hall Institute of Medical Research in Melbourne and a board member at our parent organization, told us he believes this study is an important step towards a better scientific record:

The paper from Bik, Casadevall and Fang provides strong evidence supporting the idea that a major reason for the lack of reproducibility of science papers is deliberate falsification of data.

They completed the Herculean task of visually inspecting the figures in over 20 thousand papers, looking for image duplications. They then looked at suspicious images more closely using image processing software. In this way they detected duplicated images in 3.8% of papers. Although they could not distinguish accidental duplications (e.g. by incorporating the same image file twice in a multi-part figure), they did sub-categorize the papers into less worrying classes such as “cuts”, “beautification”, and “simple duplications” and more worrying classes in which the duplicated images were repositioned or further altered such as by stretching or rotating.

Although they mainly focused on papers with Western blots, their findings in no way suggest that Western blotting is a flawed method.

Indeed, it suggests that Western blots are harder to fake in an undetectable way than other experimental data.

Retraction Watch
The strength of their evidence should be enough to convince everyone that there is a major problem with how research is being conducted. Now we need to determine what to do with this information. Should journal implement similar visual screens, should they use computerized screens, or some sort of combination? What is the best way to handle the suspicious images are detected?

Bik told us she believes peer reviewers and editors should be more aware of inappropriate duplications, and take the time to look for this kind of problem before a paper gets published, so journals don't have to deal with retractions after the fact.

"You should be flagging this when you see it, give authors a chance to do something" like fix the questionable figures, she said. "It's better to reprimand somebody in private, rather than in public."

Like Retraction Watch? Consider making a tax-deductible contribution to support our growth. You can also follow us on Twitter, like us on Facebook, add us to your RSS reader, sign up on our homepage for an email every time there's a new post, or subscribe to our new daily digest. Click here to review our Comments Policy. For a sneak peek at what we're working on, click here.

SHARE THIS:



RELATED

Author denies accusations of blatant duplication

June 10, 2016

In "author objections"

Authors retract two plant biology papers over duplicated images

June 5, 2017

In "duplication retractions"

Plant biologist earns string of retractions, bringing total to 9

April 24, 2017

In "biomed research intl"

📅 April 19, 2016 👤 Cat Ferguson 📁 duplication retractions, image manipulation, united states

16 thoughts on “One in 25 papers contains inappropriately duplicated images, screen finds”



Enrico Bucci

April 21, 2016 at 6:06 am

Only 4%? This is at odd with what i see



Morty

April 21, 2016 at 7:52 am

I agree with Bucci. When screening for duplications of all figures, I get much higher % duplications. This is in cancer research journals.



Monya Baker

April 21, 2016 at 5:36 pm

If you look at the 10 cancer journals among the 40 covered, the average rate of duplications. Still, too small a set to make

Follow 2 have higher than

Follow this blog



Phil Davis

April 21, 2016 at 8:10 am

I don't understand Bik's claim that retracting an article is a Factor. As explained repeatedly on this blog, journal editors should have people trained to investigate the issue, follow up with authors and their institutions, and more importantly, be able to defend themselves against threats of lawsuits.

Get every new post delivered right to your inbox.

Email address



TL

April 21, 2016 at 9:23 am

The whole “exposition of scientific results by appeal to Western blots” paradigm seems to be coming to a crashing end.

By the way, surely a mechanical image-recognition task like this is a simple job for a computer vision algorithm? Maybe in the future blots like this will need to be submitted to TurnYourBlotIn for automated checking at submission time?



fernandopessoa

April 21, 2016 at 10:41 am

“By the way, surely a mechanical image-recognition task like this is a simple job for a computer vision algorithm?” That’s what the chattering classes think.



Conrad Seitz MD

April 21, 2016 at 9:40 am

What surprises me is that these image duplications weren’t caught by the peer reviewers at the time the papers were submitted. The example shown is obvious, and likely many more were equally obvious. You don’t need a trained eye because your eye has already been trained– to recognize faces. Why weren’t these caught in peer review? Possibly the review was superficial at best, negligent at worst.



Helene Z Hill, PhD

April 21, 2016 at 10:04 am

Clare Francis is a whiz a spotting such things. Haven’t heard from him/her for some time. Hope s/he is still at it. And what about analysis of numbers — terminal digits and the like. Someone needs to get on to that.



Ken Pimple

April 21, 2016 at 12:08 pm

The Office of Research Integrity has a number of “forensic tools” at <http://ori.hhs.gov/forensic-tools>



Jaime A. Teixeira da Silva

April 21, 2016 at 11:34 am

It’s encouraging to see this effort by Bik. My hat off to her.

In plant science (pure and applied), my personal estimates of any “irregularity” in papers suggests that the percentage of image manipulation, including partial or full duplication, band and gel manipulation, far exceeds the 4% that Bik et al. claim. A crude estimate sits at 5-10% of the examined literature, with a rough total of about 6000 papers examined, using techniques not all that different to those used by Bik’s. Also, similar to Bik, three plant science journals rejected a PPPR report on a specialized crop species, for similar reasons: scope, not in the interests of readers, etc. But there is hope.

Very unfortunately, unlike what Phil Davis claims above, the percentage of COPE member (journal or publisher) editors who are not responding to named reports (I, also like Bik, largely abandoned the anonymous / pseudonymous mask in early 2016), not even sending an acknowledgement email, and not following up on requests for an inquiry or investigation is also in double digits (again, for plant science-related queries). There are a few laudable exceptions, however, which also bring hope, and thus also place pressure on members and non-members alike, to respond.

In my experience, one way of making sure that a public record exists is to post to a “centralized” page like at RW [1], while also incorporating PubPeer entries into any formal report [2]. It’s a slow and painful process to dredge through so much literature, but an essential step in trying to heal a serious situation (again, in the plant science literature).

[1] <http://retractionwatch.com/2014/01/25/weekend-reads-trying-unsuccessfully-to-correct-the-scientific-record-drug-company-funding-and-research/#comments>

[2] Teixeira da Silva, J.A. (2015) A PPPR road-map for the plant sciences: cementing a road-worthy action plan. Journal of Educational and Social Research 5(2): 15-21.

<http://www.mcser.org/journal/index.php/jesr/article/view/6551>

DOI: 10.5901/jesr.2015.v5n2p15



e-Patient Dave

April 21, 2016 at 11:48 am

Did any of these studies analyze whether open peer review has a better or worse rate of missing these incidents?



Laughing Man

April 21, 2016 at 12:51 pm

Even Graphs are manipulated.

**John Krueger**

April 22, 2016 at 1:46 pm

Bik et. al conclude that the “prevalence of papers with problematic images rose markedly during the past decade.” The temporal change (shown their Figure 5, bioRxiv) jives with the increase in questioned images in ORI cases during that same period (see Krueger, ORI Newsletter, vol 13 no3, 2005; vol. 17, no 4, 2009; Acct in Res. 9: 105-125, 2002); Yet, while not ‘revelatory,’ Bik’s data is vastly more comprehensive!) Fair enough, but by quantifying their result relative to a “paper”, one can’t know what to make of Bik’s data. Only by knowing how many more “images” per paper were actually published during this time period can one draw any conclusions. For example, does a change in image incidence account for why there are so many “NA” results in their data for the earlier papers (see their table)?

The question remains (as I have asked elsewhere: ORI NWSLTR vol 21 no 1, 2012), as to i) whether the increased incidence of retractions tells us that researchers are becoming less trustworthy, or ii) is it simply that an increased reliance upon Imaging in Science means that research reports are more transparent? (In other words, increased reliance on image-data make the “one-percenters” easier to detect?) Neither Bik et al’s nor my approach answers that question. However, starting with an operational definition of an “image-in-science” (as the independent variable), and slogging through the literature, might provide a direct answer. If Bik et al have that data in their data sets, they should exploit it. The answer could also be relevant to the RW post by Wagers and Marusic about whether exhortative based training in research ethics is worth while. What is the best way to spend resources to address misconduct?

**Alan R Price**

April 22, 2016 at 2:37 pm

I agree with John Krueger, who began working on such image misconduct frequency tracking in the ORI cases we handled, beginning two decades ago and continuing ever since then, publishing his analyses in ORI newsletters, association speeches, and journals. His forensic droplets and other image analysis tools are still available on the ORI website

**Anonymous**

June 19, 2016 at 4:13 am

“...there is a major problem with how research is being conducted. Now we need to determine what to do with this information” – How about improving career paths and easing the pressures on researchers (at all levels) to publish (or-perish). Cure the disease, not just focus on its symptoms.



Ulrik John

March 7, 2017 at 7:59 pm

Look closely at the example of the western blot in this post: elements of the band in lane 7 are indistinguishable from the duplicates in lanes 9 and 10, and those in 6 and 8 also share distinct features in common!

Great initiative Elies, that deserves the widest possible dissemination.

Proudly powered by WordPress