The Journal of Clinical Investigation

All data are not created equal

Ushma S. Neill

Editorial

An inauspicious start to 2009, unfortunately. This issue may seem a little thinner than others we've recently published, as 4 articles that were previously accepted and scheduled for publication in this issue will not appear. We continue to screen all figures from accepted manuscripts, and we continue to find irregularities. In several cases, the alterations in the figures led to the discovery of some fundamental problems with the data. Many of the papers suffered from the same problems, and this led us to consider whether it was time to revisit some experimental basics.

Find the latest version:

https://jci.me/38802/pdf

All data are not created equal

An inauspicious start to 2009, unfortunately. This issue may seem a little thinner than others we've recently published, as 4 articles that were previously accepted and scheduled for publication in this issue will not appear. We continue to screen all figures from accepted manuscripts, and we continue to find irregularities. In several cases, the alterations in the figures led to the discovery of some fundamental problems with the data. Many of the papers suffered from the same problems, and this led us to consider whether it was time to revisit some experimental basics.

Shall we start with some basic biochemistry? For the examples in this editorial, let's discuss Western blots in particular (though the rules apply to Northern, Southern, and PCR blots too). We've stated before (1) that bands can be spliced together, but only, and I repeat *only*, if they were noncontiguous but run on the same gel at the same time. The figure then needs a thin line in between the spliced lanes and appropriate text added to the figure legend to reflect the modification. Also, there is nothing more reliable than a blot with bubbles — there is no need to erase background noise or doublet bands.

I was always taught that a loading control, say β-actin or GAPDH, was probed off the same gel. My understanding of a loading control is that it represents an analysis of an irrelevant protein from the exact same gel lane to assess how much sample was loaded in that particular lane. This seemed to be an elementary governing principle to me, which was why I was somewhat shocked when we had an article that, among other problems, ran the loading control on a separate gel at the same time. To me, running a parallel gel, even if the sample run on the gel was an aliquot from the same tube, does not demonstrate equal loading of sample in the experimental gel. How does one control for variations in pipetting such small volumes? I realize that some blots need to be stripped and reprobed several times, but can't you just cut off the bottom and reprobe that part for the actin or other loading control and reprobe the top for another protein of interest? Or can't you run duplicate blots and present 2 rows of loading controls? We are worried about the manipulations we can detect, and I suspect that in cases such as these, a heavy (or light) hand with a pipette can influence a band's appearance when no loading control is there to normalize it.

Another basic principle relates to exposure times. Specifically, in a particular row of a Western, all bands presented should be from the same exposure time of the film. We allow band splicing — with all the appropriate caveats, as described above. But splicing lanes from various exposures of the same blot doesn't prove anything. If you expose most any gel long enough, you'll get bands to appear and the results you want. Space is not limiting, and you can always add another row of lanes to your figure that show a longer exposure time to verify that the protein was there, albeit in trace amounts.

This may seem elementary, but *keep the raw data*. A well-annotated lab notebook can resolve problems very quickly. Scan the film from your blot and save it on your computer and on the lab server, and keep a copy on a USB drive or elsewhere (anecdotally, we once had an author claim he could not provide a high resolution version of a Western blot from his accepted paper because all the data were on a single USB drive that had been lost). Then, paste the film into your lab notebook. Label the film clearly and annotate the date and conditions. And protect the lab notebook. Do not lose it when you move your lab (another excuse from an author when he could not substantiate the data in his accepted manuscript). Perhaps even e-mail a scan of the uncut, labeled film to yourself and the senior author to ensure the data are accessible. And further than just saving immunoblots — the same relates to histology and other data — print them out. Make multiple copies. Label the data carefully. Keep clear records.

When questioned about data in a paper, many senior authors feign ignorance and blame the first author or data generator, but I find that disingenuous. If you are the senior author, it is incumbent on you to verify all of the raw data yourself. There is intense pressure to produce, and to produce high-impact results. Sometimes this can lead to a student doing anything to please. If the paper goes out with your name on it, you should be able to verify every single piece of data in it and take responsibility for it.

We have tools to be able to detect whether you have altered your figures in any way. If you cut it, crop it, squeeze it, tease it, or otherwise massage it — we can see it. And you can be sure that we have closely examined all figures in the 19 other articles in this issue. I hope that the experiments themselves in our published papers have been performed properly, but that is not something we can police.

Ushma S. Neill
Executive Editor

OK, easy now. Just a little nip here and there. We don't want it to look like it's had any work done.

1. Neill, U., and Turka, L.A. 2007. Navigating through the gray (and CMYK) areas of figure manipulation: rules at the JCI. *J. Clin. Invest.* **117**:2736.