Mike Rossner
Comments on STM "Recommendations for handling image integrity issues"

I applaud the STM for working to codify principles for editors to follow when handling image data integrity issues. Here are a few specific comments about the document:

1. The document does not address how to accomplish the important step of image screening. This has been a contentious issue for nearly 20 years, and the working group may have been avoiding it. In my opinion, journals have an obligation to screen all images in all figures in all manuscripts accepted for publication. There are various ways to accomplish this systematic, universal screening, such as, visual screening in-house, visual screening outsourced, and algorithmic methods that are now coming online and need to be vetted for effectiveness by comparing them to visual screening. STM might consider creating recommendations for the screening process—what should be screened and how it can be screened.

2. The phrase, **"images and data"** appears several times in the document. In my opinion, this undermines the notion that images ARE data (see this article by Doug Cromey from 2014). The phrase "image data" would be a better choice.

3. Principle (2), paragraph 3. **"Authors should be informed in advance if the editors plan to approach the corresponding authors' institution."** I disagree with this principle, and I think many Research Integrity Officers may also disagree with it. Institutions often need to have the opportunity to sequester data before an inquiry. Informing the author before informing the institution might give the authors a chance to conceal data before they are sequestered.

4. Principle (3), paragraph 1. **"Source data in this context is defined as minimally processed "raw" data underlying a figure."** I think it should be defined as unprocessed, "raw" data, and this definition might be expanded to specify, for example:

   > For blots, this constitutes the actual piece of X-ray film, or, if the data were acquired digitally, the scan of the complete blot in the file format in which it was originally saved (either TIFF, which is a lossless format, or preferably the proprietary format generated by the imaging system). Images that have been cropped and/or imported into another application to compose a labelled figure do not constitute source data.

5. Principle (7), paragraph 1. **"Editors may decide not to pursue non-definitive issues and minor issues that would not affect the main conclusions in a published research paper."** In my opinion, a question raised about any piece of data in a published article is important. It may lead to an editorial expression of concern about, or a correction of that particular piece of data (even if the main conclusions of the paper are unaffected), which is important information for the reader.

6. Principle (8): **"In some cases…"** Some examples would be helpful here.

7. In my opinion, the classification system as presented in the table does not represent a logical approach to handling image integrity issues. A more logical approach would be a decision tree, because the initial investigative actions by a journal editor are the same regardless of the "severity" of the manipulation. Here's a very rough example:

    A. Is there reason to suspect that an image has been manipulated? If "no", proceed to publication. If "yes", request the source data from the authors.

    B. Do the authors have the source data? If "no", see Principle (3). If "yes", go to "C".

    C. Are the source data accurately represented in the submitted figure (i.e., the figure is not manipulated)? If "yes", proceed to publication. If "no", proceed to "D".

    D. Does the manipulation affect the interpretation of the data? The "no" answer effectively corresponds to Level I in this document, and the "yes" answer effectively corresponds to Level II in this document.

    E. Recommendations presented in the table provide further branches to the decision tree, such as, Do we withdraw the submission? Do we publish an expression of concern? Do we publish a correction? Do we publish a retraction? Do we report the manipulation to the corresponding author's institution?

    Level III in this document relates to "intent", and I do not think that intent is relevant to the decisions that a journal editor has to make (publish/not publish a manuscript submitted to the journal, or publish an expression of concern/correction/retraction of a published article). It certainly is relevant to the decisions that an institution has to make about the determination of misconduct and the consequences of that determination.