Current Biology, Volume *24*

Supplemental Information

The Availability of Research Data Declines Rapidly with Article Age

Timothy H. Vines, Arianne Y.K. Albert, Rose L. Andrew, Florence Débarre, Dan G. Bock, Michelle T. Franklin, Kimberly J. Gilbert, Jean-Sébastien Moore, Sébastien Renaut, and Diana J. Rennison

Supplemental Experimental Procedures

This supplemental file contains the text of the email messages sent to the authors of the 516 papers included in our survey.

When authors replied asking for more information, we provided additional details as required.

1) **Initial email** (sent April 15th)

To: Corresp, First, Last
Subject: Request for data from your [year] paper in [journal]

Dear Drs. X, Y, and Z (first author, corresponding author, and senior author),

We are currently conducting a study to determine how reproducibility and data availability change through time. We are focusing on the reproducibility of discriminant function analyses of morphological data, and so we would be delighted if you could:

- send us the morphology data used in your [year] paper [title] published in [publishing journal].
- tell us how long it took to find the data: .... minutes

If you are unable to send us the data, we would very much like:
- an estimate of how long it would take to get the data: .... minutes
- to understand why the data are unavailable (e.g. dataset lost, stored in an inaccessible format, required for ongoing study)

Our analysis will not identify individual studies, but will instead focus on overall patterns and trends. Your responses will therefore be completely confidential.

Many thanks for your time and help. Please do not hesitate to get in touch if you have any questions.

Sincerely,

The UBC Reproducibility Group
http://www.zoology.ubc.ca/~repro/UBC_Reproducibility_Group

2) **Follow-up email** (sent May 8th 2013 if no response to our initial email was received).

To: Corresp, First, Last
CC: additional emails

Subject: Request for data from your [year] paper in [journal]

Dear Drs. X, Y, and Z (first author, corresponding author, and senior author),

We're following up on an email we sent three weeks ago: we are currently conducting a study to determine how reproducibility and data availability change through time. In particular, we are focusing on the reproducibility of discriminant function analyses of morphological data, and so we'd be delighted if you could:

- send us the morphology data used in your [year] paper [title] published in [journal]

- tell us how long it took to find the data files : .... minutes

If you are unable to send us the data, we'd very much like:

- an estimate of how long it would take to find the data files: .... minutes

- to understand why the data are unavailable (e.g. dataset lost, stored in an inaccessible format, required for ongoing study):

Our analyses will focus on overall patterns and trends, and papers and authors will not be identified in the study. Your responses will therefore be completely confidential. We've included a more detailed cover letter below, but please don't hesitate to get in touch if you have any questions. Many thanks for your time and help.

Sincerely,

Tim Vines
on behalf of the UBC Reproducibility Group

To whom it may concern:

We are a group of researchers at the University of British Columbia that studies issues related to data archiving in scientific research, and we're currently examining how the availability of research data changes with time since publication. For example, data can be lost when researchers leave science and can no longer be contacted, or when data are stored on outdated electronic media. In addition, we also want to test whether time since

publication affects the reproducibility of statistical analyses, perhaps because standards for describing analyses have changed over time, or because the exact data used in the analysis are no longer available.

To quantify these phenomena, we are requesting data from papers that used a discriminant function analysis (DFA) on morphological data. The methods used to collect morphological data (e.g. calipers and microscopes) and the statistical underpinnings of DFA have remained constant for many years. We randomly selected a subset of these studies published in odd years between 1991 to 2011, and we are currently asking for the data underlying the DFA directly from the authors.

For datasets that we do receive, we will attempt to reproduce the results found in the paper, following the methods outlined by the authors. We have discussed our study with the UBC Behavioural Research Ethics Board, and they determined that it did not need to go through the ethical review process. Since the financial resources required are minimal we have not applied for external funding.

Please do not hesitate to contact us with questions, or also see the previous publications from our group by pasting the following link into your web browser: http://www.zoology.ubc.ca/~repro/UBC_Reproducibility_Group/Welcome.html

Sincerely,

The UBC Reproducibility Group

Dr. Rose Andrew
Dan Bock, PhD student
Dr. Florence Débarre
Dr. Michelle Franklin
Kimberly Gilbert, PhD student
Dr. J.-S. Moore
Dr. Sébastien Renaut
Diana Rennison, PhD student
Dr. Tim Vines