Did awarding badges increase data sharing at *BMJ Open*? A randomised controlled trial

Rowhani-Farid, Anisa[1,2]; Aldcroft, Adrian[3]; Barnett, Adrian[2]
[1] University of Maryland, Baltimore
[2] Queensland University of Technology
[3] BMJ Publishing Group

INTRODUCTION

- Sharing data and code are important components of reproducible research
- Data sharing in research is widely discussed in the literature, however there are seemingly no evidence-based incentives that reward data sharing, nor randomised studies that demonstrate the effectiveness of data sharing policies at increasing data sharing
- A simple incentive, such as an open data badge, might provide the change needed to increase data sharing in health and medical research

METHODS

1. Parallel group RCT
2. Participant recruitment, CONSORT flowchart

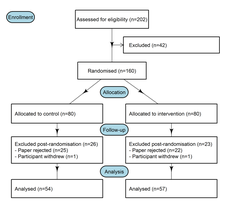

Figure 1: CONSORT flowchart for the RCT testing badges at *BMJ Open*

3. Statistical analyses
   - Primary outcome: **data sharing rate**
   - Fisher's exact test as there were small cell sizes
   - Calculate the percent shared in each treatment arm, the difference between the arms, and a 95% confidence interval of the difference

RESULTS

Table 1: Numbers of authors receiving a badge by treatment group – control and intervention

Awarded a badge	control n	%	intervention n	%
Yes	2	4	2	4
No	52	96	55	96
All	54	100	57	100

- The odds ratio for awarding badges in the intervention group relative to the control is **0.9 with a 95% confidence interval from 0.1 to 9.0**. The **p-value from the Fisher's exact test is 1**.

DISCUSSION

- Badges did not motivate researchers that publish in *BMJ Open* to publicly share their raw data as the odds ratio of awarding badges in intervention group relative to the control is close to 1 (0.9). However, given that the confidence interval is wide (0.1 to 0.9), we cannot completely negate a possible badge effect.
- There was no difference in the data sharing statements, ranking of top words used in data sharing statements, and in number of words in statements between control and intervention groups

BMJ Open

QUT

Open data **badges did not motivate** researchers who published in *BMJ Open* to **publicly share their raw data**

Image source: Open Science Framework, 2016

Protocol

Contact details

Table 2: Numbers of final data sharing statements by treatment group – control and intervention

Final Data Sharing Statement	control n	%	intervention n	%
No additional data is available	21	39	23	40
Data is available upon request	30	56	32	56
Data is available at a third party depository	3	6	2	4
All	54	100	57	100

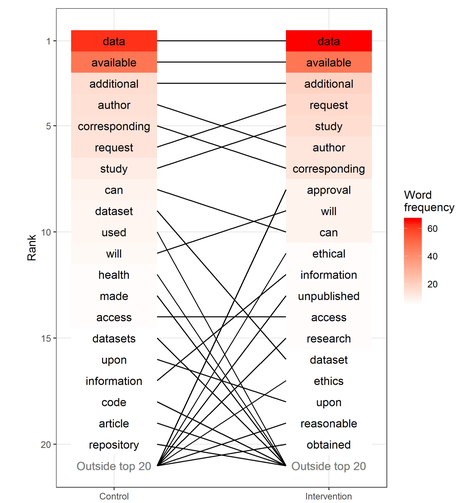

Figure 2: Plot of difference in ranking in the top 20 words between treatment group – control and intervention

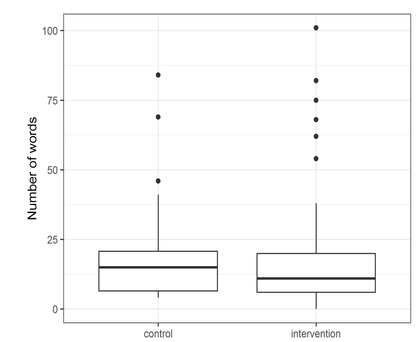

Figure 3: Box plot of the number of words used in the final data sharing statements by treatment group – control and intervention

Poisson regression model of word counts in data sharing statements: there was no association between the intervention and the number of words. The mean rate ratio was 1.05 with a 95% CI from 0.73 to 1.53.

Time needed to check for open data:
Mean time needed to check for open data by QUT study team was 3 minutes (n=8, min=1 max=5 mins)