Association of Learned and Professional Society Publishers

Introducing a data availability policy for journals at IOP Publishing: Measuring the impact on authors and editorial teams

Jade Holt [1], Andrew Walker [2], and Phill Jones [3]*

[1]IOP Publishing, Bristol, UK

[2]TaskAdept, Basingstoke, UK

[3]MoreBrains Cooperative, Edinburgh, UK

ORCID:
J. Holt: 0000-0002-0227-1950
A. Walker: 0000-0002-6036-7918
P. Jones: 0000-0003-0525-6323

***Corresponding author:** Phill Jones, MoreBrains Cooperative, Edinburgh, UK.
E-mail: phill@morebrains.coop

Abstract

As the open research movement continues to gather pace, a number of publishers, funders, and institutions are mandating the sharing of underlying research data. At the same time, concerns about introducing extra quality control steps around data availability statements (DAS) are driving a discussion about the best way to make data more open without slowing down publication. This article describes a pilot project to introduce a new Open Data policy to three IOP Publishing (IOPP) journals as part of IOPP's commitment to increasing transparency and support for open science. An investigation was undertaken using an automated workflow monitoring tool to understand the impact of this change on authors and the editorial staff. Changes in revised submission processing times and how often manuscripts were returned to the author were measured. An overall increase in the time editorial staff spent processing manuscripts was found as well as an increase in the number of times manuscripts were returned to authors. Detailed analysis shows that manuscripts in which authors claim in the DAS to have included data within the manuscript were the most strongly affected. Steps to mitigate the effects through improved author communication were found to be effective.

Keywords: DAS, Data, data availability, open data, time-tracking

INTRODUCTION

Increasingly, research funding agencies, institutions, and scholarly publishers are introducing policies, processes, and mandates in areas like open access (Piwowar et al., 2018), open research (Howard Hughes Medical Institute, 2017; Hrynaszkiewicz et al., 2020), reproducibility (AEA, 2019), Daugherty (Alan et al., 2016; Science et al., 2019), and data management (e.g., FAIR Data) (Directorate-General for Research & Innovation, 2016). There is growing evidence for the positive benefits of increased openness to improve research reproducibility (Colavizza et al., 2019; Piwowar & Vision, 2013) and robustness (Munafò et al., 2017).

Meanwhile, there is also increasing interest in shortening the time it takes for research to have a real-world impact (Amat, 2008; Hanney et al., 2015; Powell, 2016). Some authors have suggested recent developments with the COVID-19 pandemic, and Zika before it, illustrate the need to more rapidly disseminate research findings (Kupferschmidt, 2020).

When taken together, these two drivers represent a potential tension between increased quality control requirements and more rapid dissemination. Indeed, as Federer et al. stated in their 2018 article on data sharing at PLOS (Federer et al., 2018), there was some concern that PLOS' progressive stance on data sharing mandates might dissuade authors from publishing with them. This tension is not a new issue for scholarly publishers and has in the past been addressed by creating a range of publication types. For instance, *Rapid Communications* and *Letters* journals are intended to have quicker turn-around times. On the other hand, with research becoming increasingly global and interdisciplinary (Digital Science et al., 2017), overlapping funder, institutional, and community requirements are making publishing workflows increasingly complex and challenging to manage. The need for efficient editorial and production workflows is therefore more important than ever for scholarly publishers to meet the needs of their stakeholders.

Data Availability Statements (DAS) are sections of a research article that inform readers where the underlying research data associated with an article is, and under what conditions it can be accessed or reused. The Transparency and Openness Promotion (TOP) Guidelines from the Centre for Open Science describe three levels of data transparency that map to requirements for DAS that publishers might implement (Nosek et al., 2014). The lowest of these levels merely asks researchers to include a statement indicating if data can be shared; the most stringent requires public availability of data (with exceptions), and independent verification of analyses.

In July 2019, IOPP announced a new data availability policy to be piloted on three journals; *Environmental Research Letters (ERL)*, *Journal of Physics: Complexity (J. Phys. Complex)*, and *Machine Learning: Science and Technology (MLST)* (IOPScience, 2017). The policy requires all accepted articles to include a DAS stating whether the data are accessible plus a persistent identifier linking to the data location and licensing terms when appropriate. The goal of the pilot was to assess the effects of requiring a DAS as part of the publishing workflow. Specifically, IOPP were interested in measuring any increase in manuscript processing time and the number of times articles were returned to authors to be corrected for non-compliance with the mandate or other quality control reasons. We define the rate of return to authors as the unsubmit rate. This is not to be confused with the rate at which authors choose to withdraw manuscripts from consideration.

It was hypothesised that the introduction of the DAS mandate would increase the total processing time for manuscripts; and to measure the effects of workflow changes, IOPP worked with TaskAdept (www.taskadept.com) as a development partner on their workflow intelligence solution, Carina.

As part of the pilot, IOPP partnered with figshare (www.figshare.com) for seamless repository deposition.

Pilot design and methodology

The Carina software from TaskAdept captures detailed timing data for tasks in browser-based workflow platforms. In this study, it was applied to editorial and production staff workflows when processing journal submissions. Editorial workflow events (submission checks, reviewer selection, *etc.*) were timestamped automatically using a browser plugin without the use of manual time tracking. Data were automatically aggregated to reveal insights into workflow resource use, bottlenecks, and potential inefficiencies.

Of specific interest was the total time spent by editorial staff on verifying submissions for completeness based on a standard checklist for each manuscript that included compliance with the DAS policy. The total submission processing time for each manuscript was calculated as the sum of the time spent for all revised submission checklist tasks. This included time spent on successfully completed checks and on unsuccessful checks that resulted in the manuscript being unsubmitted back to the author. The average number of unsubmissions for a group of manuscripts is defined as the *unsubmit rate*.

Submission processing times and unsubmit rates for all three journals in the pilot study were monitored over a period of 4 months (March–June 2019) to establish baseline values. *J. Phys. Complex* and *MLST* are new journals. At time of writing, *MLST* had published its first issue in March 2020, while *J. Phys. Complex* published its first issue a few months later in June 2020. Naturally, copy flow for new journals is variable and hopefully increases over time. Inconsistencies in copy flow would create serious confounds in the analysis. For this reason, only data from *ERL* is used for analysis here. This provided a baseline dataset for 272 manuscripts submitted over this period.

From 1st July 2019, instructions for authors included guidance on the inclusion of a DAS for all manuscripts. DAS verification was not part of initial editorial quality checks upon first submission. The decision was taken not to include checks at this point so as not to overly burden editorial staff or authors of

---

Key points

- Introduction of extra quality control steps like a data availability statement (DAS) can result in extra time and cost to the editorial process.

- Introducing extra quality control steps can also increase the frequency of quality control failures leading to more manuscripts being sent back to for amendment and resubmission.

- Increases in editorial and author burden varies by the type of DAS, with those where the author claims data is available in the manuscript being the most burdensome.

- Increases in editorial and author burden can be reduced substantially, if not eliminated, through clear and timely communication with authors.

- The use of automated time-tracking software enables quantitative comparisons of time and cost changes as a result of amendments in editorial workflows.

TABLE 1 DAS levels and descriptions.

Level	Descriptions
Openly available	The data that support the findings of this study are openly available.
Following delay	The data that support the findings of this study will be openly available following a delay after publication.
Upon request	The data that support the findings of this study are available from the corresponding author upon reasonable request.
Within manuscript	Any data that support the findings of this study are included within the Manuscript.
Not applicable	Data sharing is not applicable to this article as no new data were created or analysed in this study.

papers that would not go on to be accepted for publication. Authors were reminded of the need for the DAS in decision letters when asked to make minor or major revisions. The reminder read:

> A mandatory data availability statement policy applies for this journal. Please ensure a data availability statement is included in your manuscript. Instructions for including a data availability statement are available at https://publishingsupport.iopscience.iop.org/questions/data-availability-statement-policy/

A further reminder was included in the resubmission form for corrected manuscripts. DAS verification was included as part of quality checks for manuscripts returned after revisions.

Guidelines for authors included a range of model statements to help authors write an appropriate DAS. Table 1 shows the suggested levels.

The distribution of total measured processing times across all revised manuscripts during the baseline phase was observed to be non-normal, necessitating the use of non-parametric statistical analysis. Average total times for submission processing was calculated by median. Median unsubmit rates were also calculated. Kruskal-Wallis rank sum tests were used to test for the significance of differences between phases of the trial. The identification of which specific groups were significantly different from baseline values was done using Wilcoxon rank sum test using Holm correction for multiple comparisons.

The initial trial lasted until August 31st 2019. After the initial impact of changes to journal workflows were measured, it was decided that increases in article processing times and unsubmit rates were greater than desired. The pilot study was therefore extended and targeted workflow changes were introduced to mitigate the negative impact on time to publication and number of unsubmissions.

Instructions to authors regarding the DAS were clarified and made more frequent. Improvements were three-fold.

1. In decision letters, information about the DAS requirements were included in an attachment titled *DATA-AVAILABILITY-STATEMENT-INCLUSION---PLEASE-READ.pdf*.
2. The following text was added to existing revision deadline reminder emails: *Please include one of five data availability statements in your paper. For more information about data availability statements, visit our website Publishing Support* https://publishingsupport.iopscience.iop.org/questions/data-availability-statement-policy/.
3. A new dedicated email was sent automatically between the first decision and revision deadline that outlined the policy and contained the five model data statements. Data continued to be gathered during this mitigation phase until October 31st 2019.

RESULTS

A summary of the data is shown in Table 2. Upon initial inspection, it can be seen that the average processing times are higher during the initial trial phase compared to baseline, although care must be taken when interpreting results for the *Following delay* DAS type as the number of manuscripts is low at only 3 during the initial phase and 6 during mitigation. The median unsubmit rate for Baseline is zero meaning that fewer than half were unsubmitted at all (132 out of 272). This suggests that manuscript requirements were generally well understood by most authors prior to the introduction of the DAS requirement.

Kruskal-Wallis tests were applied to test for significance in the differences between average total processing time and unsubmission rates between the baseline, initial, and mitigation phases. In both cases, there was found to be an effect. In the case of total processing time, $p = 0.0002$, which is highly significant, indicating that the introduction of the DAS requirements and associated workflow changes had an effect on the amount of time it took editorial staff to complete checklists. For the unsubmit rate, $p = 0.003$, which is very significant, indicating that the DAS requirements had an impact on the likelihood and frequency with which manuscripts were returned to authors.

Total processing time

The box-whisker plots in Fig. 1 give an indication of the distributions in the total processing times for each test group. The *long tails* indicate that while most manuscripts are processed relatively quickly, with total times around 5 or so minutes, a subset required significantly more work with a small number taking as long as 30 min or more. For example, for the baseline data, 18% take longer than 10 min, with only 3.6% taking longer than 20.

There appears to be a change in distribution in average total time between baseline and the initial phase of the trial. While there is only a relatively modest increase in median total times, the upper quartiles are extended upwards, suggesting that the

Table 2 Numbers of manuscripts, average total processing times, and unsubmit rates for baseline, initial trial, and mitigation phase groups.

Phase	Author-selected DAS	Number of manuscripts	Median time	Median unsubmit rate	% Never unsubmitted
Baseline	Not requested	272	5 m 16 s	0.000	51.5
Initial	All	142	7 m 40 s	1.000	34.5
Initial	Openly available	42	7 m 47 s	1.000	40.5
Initial	Following delay	3	4 m 54 s	1.000	0.0
Initial	Upon request	55	7 m 14 s	1.000	38.2
Initial	Within manuscript	26	10 m 50 s	2.000	23.1
Initial	Not applicable	16	8 m 24 s	1.000	31.2
Mitigation	All	117	6 m 05 s	1.000	44.4
Mitigation	Openly available	43	6 m 59 s	0.000	51.2
Mitigation	Following delay	6	4 m 34 s	0.500	50.0
Mitigation	Upon request	47	6 m 05 s	1.000	40.4
Mitigation	Within manuscript	15	5 m 35 s	1.000	46.7
Mitigation	Not applicable	6	5 m 46 s	1.000	16.7

Note: In addition to results from across each phase, values for each author-selected DAS type are shown for the initial and mitigation phases. During the baseline phase, no DAS were requested or recorded.
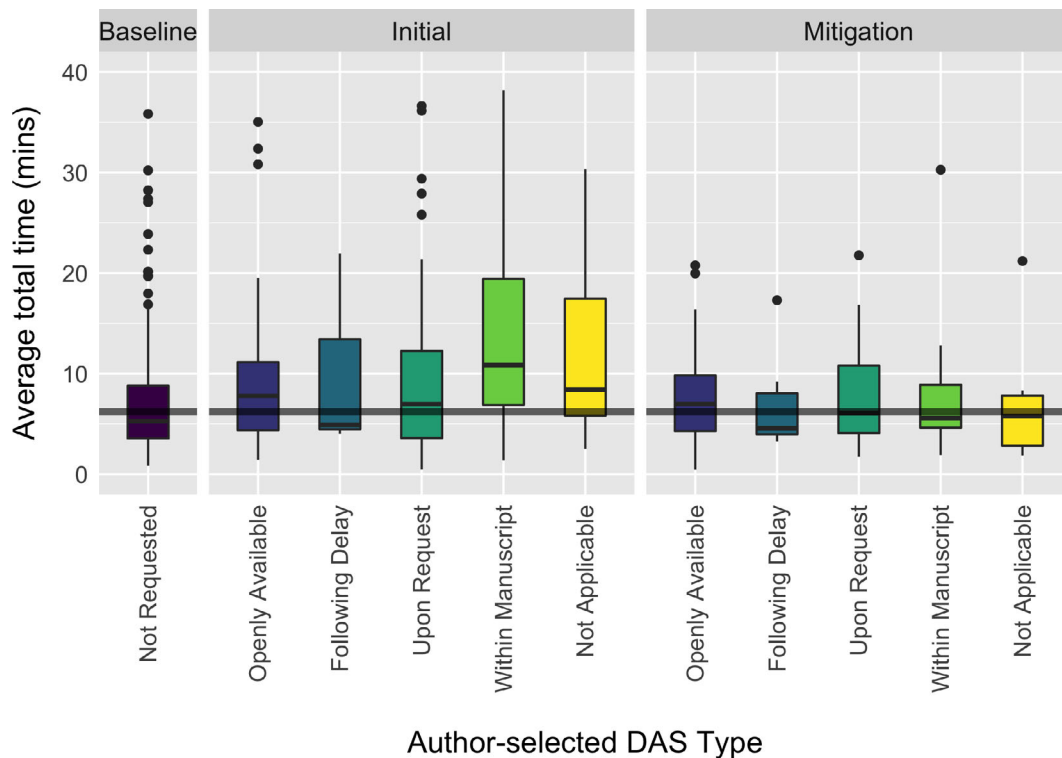

FIGURE 1 Average total time taken for editorial staff to complete manuscript checklists. For each manuscript, this includes both the time for the completed checklist added to the time taken over any partially completed checklists that resulted in the manuscript being returned to the author. During baseline, authors were not required to specify a DAS type. Separate boxes are shown for each author-selected DAS type during the initial and mitigation phases. The partially transparent horizontal line at 5 m 16 s corresponds to the median baseline processing time for easy of comparison across groups.

TABLE 3 Results of a Wilcoxon Rank test, with a Holm correction for multiple comparisons shows which groups of manuscripts had statistically significant increases in total processing time.

Phase	Author-selected DAS	n	*p*-value	Adjusted *p*-value
Initial	Openly available	42	0.323	1.000
Initial	Following delay	3	0.842	1.000
Initial	Upon request	55	0.338	1.000
Initial	Within manuscript	26	0.001	0.012
Initial	Not applicable	16	0.044	0.436
Mitigation	Openly available	43	0.684	1.000
Mitigation	Following delay	6	0.696	1.000
Mitigation	Upon request	47	0.960	1.000
Mitigation	Within manuscript	15	0.906	1.000
Mitigation	Not applicable	6	0.599	1.000

Note: Adjusted *p*-values below 0.05 are generally considered to be significant. Only the class of manuscript with 'Within Manuscript' DAS, during the initial phase showed significantly greater total processing time compared to baseline.

number of problematic articles may have gone up. Specifically, the time taken for manuscripts with the DAS type 'Within Manuscript' seems most strongly affected.

Encouragingly, the boxes based on data from the mitigation phase are shorter, suggesting fewer problematic articles.

To confirm these observations, Wilcoxon rank tests were performed for multiple comparisons between the baseline total processing times and the total processing times for each phase of the study and author-selected DAS types. The *p*-values are shown in Table 3 along with the adjusted *p*-value using Holm correction, which compensates for the increased likelihood of positive results occurring by chance when multiple sets of measurements are made.

The only statistically significant effect found was for submissions in which data was presented within the manuscripts (*p* = 0.012). The effect was present during the initial trial phase but was rescued by the mitigations. Care should be taken however when interpreting these statistics as some of the numbers of manuscripts (n) are low for certain groups. Given the pattern that appears to be shown in the box-whisker plots, it is possible that with a larger data set, more significant differences may have emerged.
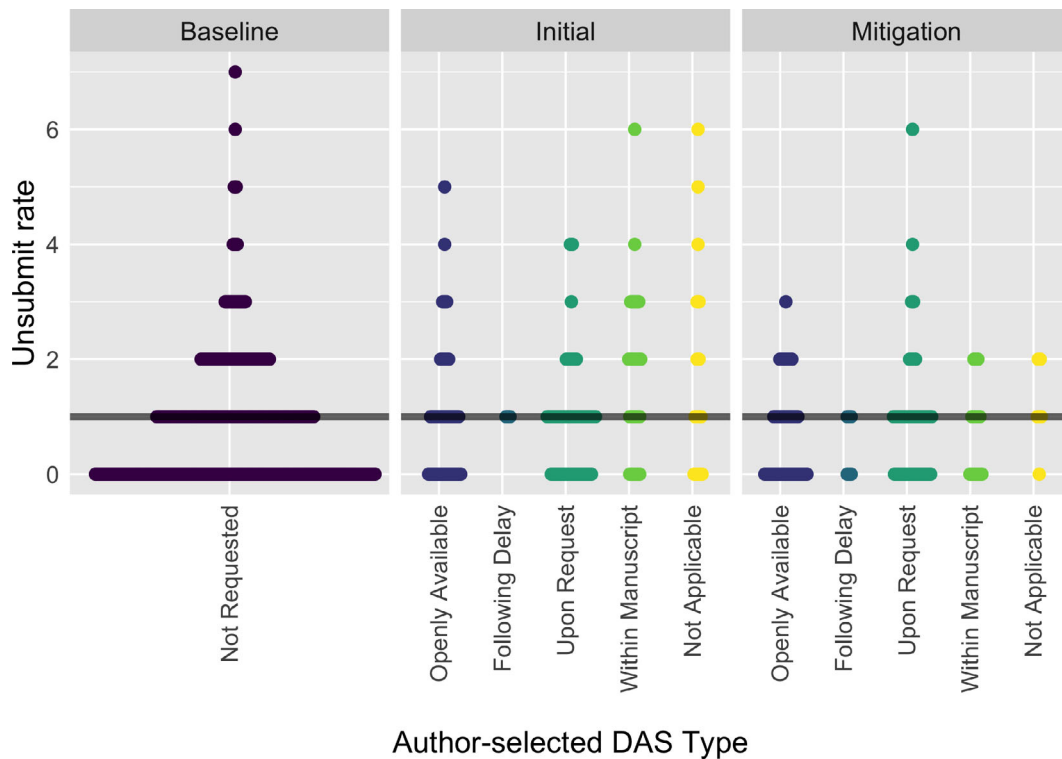

FIGURE 2 The unsubmit rates for each group of manuscripts in the study. The length of the horizontal lines represent the number of manuscripts that were unsubmitted the number of times indicated on the vertical axis. Line plots were chosen to represent these data because box-whisker plots were not informative partially due to unsubmission rates being low numbers and discrete values. During baseline, authors were not required to specify a DAS type. Separate boxes are shown for each author-selected DAS type during the initial and mitigation phases. The partially transparent horizontal line at 1 unsubmission corresponds to the median baseline unsubmission rate.
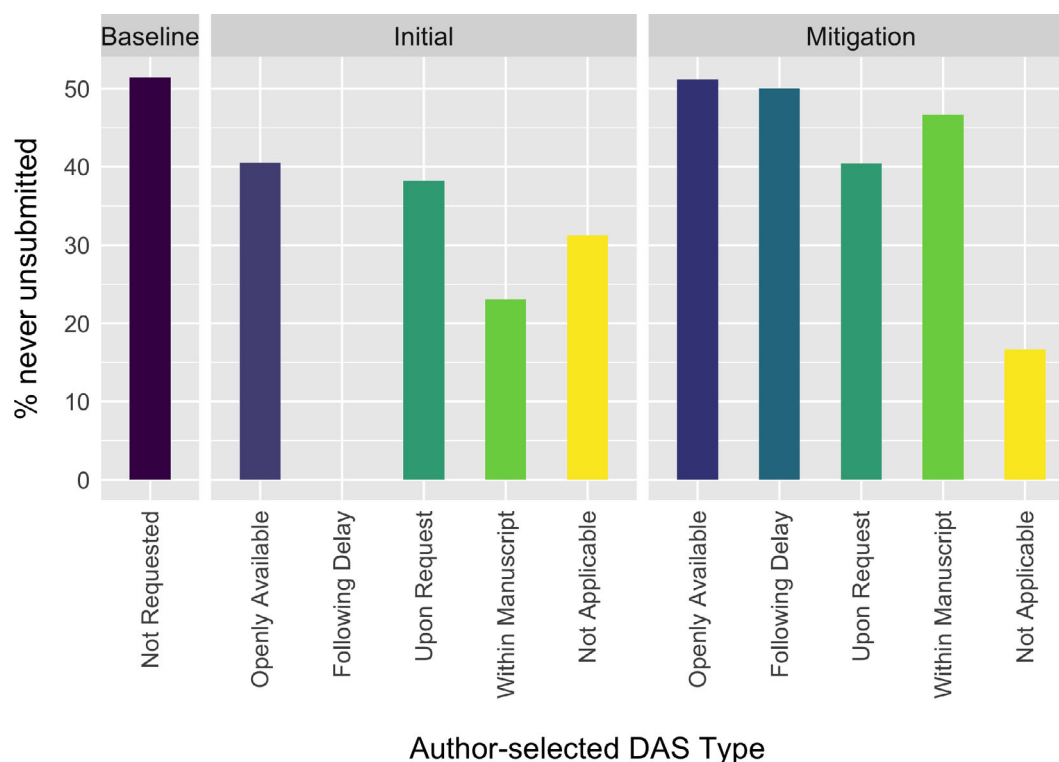
FIGURE 3 The percentage of manuscripts that were never unsubmitted for each DAS type during each phase of the trial. When a manuscript was never unsubmitted that means it passed the editorial checks on the first pass without having to be returned to the author for corrections.

Unsubmit rates

Unsubmission rates are discrete numbers and generally of the order of 1 or 2. Due to the nature of the data, box-whisker plots were uninformative. Figure 2 shows the unsubmission

TABLE 4 Results of a Wilcoxon Rank test, with a Holm correction for multiple comparisons show which groups of manuscripts had statistically significant changes in unsubmit rates.

Phase	Author-selected DAS	*p*-value	Adjusted *p*-value
Initial	Openly available	0.508	1.000
Initial	Following delay	0.444	1.000
Initial	Upon request	0.717	1.000
Initial	Within manuscript	0.001	0.008
Initial	Not applicable	0.049	0.488
Mitigation	Openly available	0.399	1.000
Mitigation	Following delay	0.488	1.000
Mitigation	Upon request	0.809	1.000
Mitigation	Within manuscript	0.808	1.000
Mitigation	Not applicable	0.234	1.000

Note: Only the class of manuscripts with 'Within Manuscript' DAS had significantly greater number to unsubmissions.

rate as horizontal line plots. The Baseline data show that prior to the introduction of the DAS mandate, most manuscripts (≈51%) have no resubmissions. The steep *inverted funnel* shape of the distribution illustrates that for the most part, manuscripts were well prepared on submission and did not require too much back and forth with authors and editorial staff.

After the introduction of the DAS mandate, Table 2 and Fig. 3 shows that the percentage of manuscripts that were successfully processed on the first attempt (i.e. were never unsubmitted) appears to fall. From inspection, it seems that introduction of the DAS mandate caused a reduction in the number of *never-unsubmitted* manuscripts and that the effect was at least partially rescued during the mitigation phase when efforts were made to make instructions clearer. Care must be taken when interpreting the data, however, as the relatively low number of manuscripts limits the statistical power of the data. In addition, the unsubmission rate, may be an insensitive measure of author and editorial staff burden.

Using the same statistical tests as were applied for total processing time, Table 4 shows that during the initial trial phase, manuscripts with the *within manuscript* DAS type could be shown to have significantly higher number of unsubmissions. It is possible that with a larger data set, significant differences would appear between other classes of manuscript and the baseline.

DISCUSSION

As the open research movement gathers pace, a number of publishers have adopted data sharing policies. The Jisc Journal Research Data Policy Registry Pilot (JRDPR) (Naughton & Kernohan, 2016), built on work carried out by Nottingham University's Centre for Research Communication from 2012 to 2104 (Sturges et al., 2015) to assess the feasibility of a registry of research data publication requirements. Based on findings from both projects, generally, policy landscapes for journals are complex, variable, and confusing for researchers. A number of publishers have developed frameworks for their own journals with specific options for open data policy types, often based on the TOP guidelines. Elsevier have developed five policies (Elsevier, 2020), while Wiley have four (Wiley, 2019). A series of two case studies from Taylor & Francis and Springer Nature (Jones et al., 2019) document the success and challenges around developing a framework for data sharing policies in which the authors observe that the various guidelines do not map well onto one another. Following on from that work (Hrynaszkiewicz et al., 2020) have developed a candidate framework for journal data sharing policies based on the TOP guidelines in hopes of unifying journal policies and reducing complexity for researchers. More recently, the STM association launched their Research Data Year (STM Association, 2020) to further the process of simplifying and unifying policy.

In light of this movement towards data sharing, it is vital to understand the time, cost, and business implications of adding extra quality control steps around open data and DAS polices to editorial processes. In particular, exploring the differential effects of different policy types will impact the differential feasibility of data sharing policies. Similar to the work that we present here, Grant at al. (Grant & Hrynaszkiewicz, 2018) measured the average additional time it took for editorial staff to process manuscripts during a trial of mandatory DAS for an initial 5, and then 20 pilot journals at Springer Nature between September 2016 and February 2017. Unlike the current study, the authors at Springer Nature used a self-reporting strategy to gather data. The average increase in processing time at Springer Nature was similar to that found here, however, the authors of that study saw their strongest increase in processing time for manuscripts where data were publicly available (≈18 m). It is possible that this difference is due to improved understanding of open science practices among authors.

The use of the Carina time tracking software allowed for automatic in-depth quantification of editorial processes without creating extra burden for staff and reducing the risk of conscious or unconscious bias. For the purposes of assessing business impact, the ability to draw out the precise amount of time that each manuscript took up during quality control checks allows for direct comparison of different DAS types. Further, precise monitoring enabled the comparison of average processing times before and after the introduction of mitigation steps to make instructions for authors clearer. In a real editorial or even publishing production environment, it is often challenging to uncover areas of opportunity to reduce overhead and waste.

There are a number of apparent patterns in the data that could not be shown to be statistically significant but are sufficient to warrant further study. For example, the average total processing times in Fig. 1 appear to show that while there is only a moderate increase in median processing time during the initial trial phase, the distributions are drawn upwards, showing that the *long tail* of the data may have increased in importance. In other words, a subset of manuscripts became more time consuming and problematic when the DAS mandate was introduced.

Increases in processing times and unsubmit rates were shown to be significant for manuscripts with the *within manuscript* DAS type. The pattern is further indicated when comparing the percentage of manuscripts that passed the editorial checks on the first try. Visual inspection of the data suggests that with a larger data set, that same trend may emerge with all of the DAS types, but it seems clear that the effects are most pronounced for the *within manuscript* DAS type. The reason for the difference between DAS types cannot be directly explained by the current study, however, anecdotal evidence from editorial staff suggests that the *within manuscript* DAS type is inherently more likely to raise red flags during quality checks. The *following delay* and *upon request* DAS types create less onerous requirements as there are fewer tasks for the author to perform prior to submission. For those authors that select the *publicly available* DAS type, it may be that they are already engaged in open research practices and therefore either already know how to share their data or have already done so.

LIMITATIONS

The DAS trial at IOPP was necessarily limited in terms of time window, number of journals and consequently number of articles. In particular, two of the journals, *MLST* and *J. Phys. Complex* were early stage with inevitably inconsistent copy flow. The number of manuscripts available for analysis was consequently limited. Particularly, the number of manuscripts where data were claimed by the authors to be available following delay (3 during initial phase and 6 during the mitigation phase) and not applicable (16 during initial phase and 6 during the mitigation phase) were too low for confident conclusions to be drawn. As a result, we have been appropriately conservative and only drawn conclusions based on data where sufficient statistical power was available.

It is likely that a larger data set would enable further statistical analysis that may verify the existence of apparent patterns in the data. For example, we suspect that a disproportionate amount of editorial time may be spent on a relatively small number of problem manuscripts that take much longer than average to process. Identifying the cause of this subset may be valuable.

During the trial, the need for the DAS was stated in the instructions to authors on the journals' websites, however, they were not checked or enforced upon initial submission. This was done to reduce the time spent on manuscripts that would go on

to not be accepted for publication. It is therefore not possible to measure any increase in processing time caused by quality checks of DAS statements prior to the point of first decision.

CONCLUSION

During the course of this study, it was found that including extra quality control requirements on journal manuscripts can increase the total processing time of those manuscripts and by extension impact costs of operation. On the other hand, careful design of workflows and communication with authors can significantly mitigate those effectsSoftware tools that embed in publishing workflows make it possible to conduct studies of the impact of publishing workflow changes without relying on self-reporting or adding further workflow burden. In this work, we made use of such an approach to measure the increase in time and by extension cost of added steps in a workflow. More importantly, such approaches enable sophisticated monitoring techniques such as A-B testing of workflows and communications with authors.

CONFLICTS OF INTEREST

Jade Holt declares that she has no conflict of interest. Andrew Walker is the owner of TaskAdept and wrote the Carina time tracking software. Phill Jones is a consultant and works with TaskAdept on a variety of projects.

ACKNOWLEDGEMENT

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available on Zenodo at https://doi.org/10.5281/zenodo.4564243 under the CC-BY-SA-4.0 licence.

REFERENCES

AEA. (2019). *AEA member announcements: updated AEA data and code availability policy*. Retrieved from www.aeaweb.org/news/member-announcements-july-16-2019

Alan, D., Hegele, R. A., Nigel, M., Rader, D. J., Marie, S. A., & Christian, W. (2016). Complying with the National Institutes of Health Guidelines and Principles for Rigor and Reproducibility. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *36*(7), 1303–1304. https://doi.org/10.1161/ATVBAHA.116.307906

Amat, C. B. (2008). Editorial and publication delay of papers submitted to 14 selected food research journals. Influence of online posting. *Scientometrics*, *74*(3), 379–389. https://doi.org/10.1007/s11192-007-1823-8

Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., McGillivray, B. (2020). The citation advantage of linking publications to research data. *PLOS ONE*, *15*(4), e0230416. http://dx.doi.org/10.1371/journal.pone.0230416.

Digital Science, Szomszor, M., & Adams, J. (2017). *Digital research report: Discovery and analysis of global research trends using GRID* (p. 7282603 Bytes). Digital Science. https://doi.org/10.6084/M9.FIGSHARE.4714249.V1

Directorate-General for Research & Innovation. (2016). *Guidelines on FAIR Data Management in Horizon 2020*. European Commission. Retrieved from https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Elsevier. (2020). *Research data guidelines*. Elsevier.com. Retrieved from https://web.archive.org/web/20200621224036/https://www.elsevier.com/authors/author-resources/research-data/data-guidelines

Federer, L. M., Belter, C. W., Joubert, D. J., Livinski, A., Lu, Y.-L., Snyders, L. N., & Thompson, H. (2018). Data sharing in PLOS ONE: An analysis of data availability statements. *PLoS One*, *13*(5), e0194768. https://doi.org/10.1371/journal.pone.0194768

Grant, R., & Hrynaszkiewicz, I. (2018). The impact on authors and editors of introducing data availability statements at nature journals. *International Journal of Digital Curation*, *13*(1), 195–203. https://doi.org/10.2218/ijdc.v13i1.614

Hanney, S. R., Castle-Clarke, S., Grant, J., Guthrie, S., Henshall, C., Mestre-Ferrandiz, J., Pistollato, M., Pollitt, A., Sussex, J., & Wooding, S. (2015). How long does biomedical research take? Studying the time taken between biomedical and health research and its translation into products, policy, and practice. *Health Research Policy and Systems*, *13*(1), 1. https://doi.org/10.1186/1478-4505-13-1

Howard Hughes Medical Institute. (2017). *Sharing-published-materials-responsibilities-of-HHMI-Authors.pdf*. Retrieved from https://hhmicdn.blob.core.windows.net/policies/Sharing-Published-Materials-Responsibilities-of-HHMI-Authors

Hrynaszkiewicz, I., Simons, N., Hussain, A., Grant, R., & Goudie, S. (2020). Developing a research data policy framework for all journals and publishers. *Data Science Journal*, *19*(1), 5. https://doi.org/10.5334/dsj-2020-005

IOPScience. (2017). *Open data policy*. IOPScience—Publishing Support. Retrieved from https://publishingsupport.iopscience.iop.org/questions/research-data/

Jones, L., Grant, R., & Hrynaszkiewicz, I. (2019). Implementing publisher policies that inform, support and encourage authors to share data: Two case studies. *Insight*, *32*(1), 11. https://doi.org/10.1629/uksg.463

Kupferschmidt, K. (2020). 'A completely new culture of doing research.' Coronavirus outbreak changes how scientists communicate. *Science|AAAS*. Retrieved from https://web.archive.org/web/20200226233716/ https://www.sciencemag.org/news/2020/02/completely-new-culture-doing-research-coronavirus-outbreak-changes-how-scientists

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Naughton, L., & Kernohan, D. (2016). Making sense of journal research data policies. *Insight*, *29*(1), 84–89. https://doi.org/10.1629/uksg.284

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Yarkoni, T. (2015). Promoting an open research

culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374.

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of open access articles. *PeerJ*, *6*, e4375. https://doi.org/10.7717/peerj.4375

Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, *1*, e175. https://doi.org/10.7717/peerj.175

Powell, K. (2016). Does it take too long to publish research? *Nature News*, *530*(7589), 148–151. https://doi.org/10.1038/530148a

Digital Science, Fane, B., Ayris, P., Hahnel, M., Hrynaszkiewicz, I., Baynes, G., & Farrell, E. (2019). The State of Open Data Report 2019 [report]. *Digital Science*. https://doi.org/10.6084/m9.figshare.9980783.v2

STM Association. (2020). *STM 2020 research data year*. STM Association. Retrieved from https://web.archive.org/web/20200608043932/https://www.stm-assoc.org/standards-technology/2020-stm-research-data-year/

Sturges, P., Bamkin, M., Anders, J. H. S., Hubbard, B., Hussain, A., & Heeley, M. (2015). Research data sharing: Developing a stakeholder-driven model for journal policies: Research data sharing. *Journal of the Association for Information Science and Technology*, *66*(12), 2445–2455. https://doi.org/10.1002/asi.23336

Wiley. (2019). *Wiley's data sharing policies*. Authoservices.Wiley.com. Retrieved from https://web.archive.org/web/20200515022613/https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/data-sharing-policy.html