of Springer Nature, through grant RG92108 "Effect of data sharing policies on articles' citation counts" granted to BM. Springer Nature provided support in the form of salaries for author IH, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing interests: One of the authors (IH) is at the time of publication in the journal, employed by PLOS, publisher of PLOS ONE. IH was employed by Springer Nature, publisher of the BMC journals, at the time of planning and conducting the research and writing of the original manuscript. This does not alter our adherence to PLOS ONE policies on sharing data and materials. There are no patents, products in development or marketed products associated with this research to declare. All other authors have declared that no other competing interests exist.

Publisher's Note: The article involves the independent analysis of data from publications in PLOS ONE. PLOS ONE staff had no knowledge or involvement in the study design, funding, execution or manuscript preparation. The evaluation and editorial decision for this manuscript have been managed by an Academic Editor independent of PLOS ONE staff, per our standard editorial process. The findings and conclusions reported in this article are strictly those of the author(s).

to give researchers more credit for sharing their work [1]. While some journals have required data sharing by researchers (authors) for more than two decades, these requirements have tended to be limited to specific types of research, such as experiments generating protein structural data [2]. It is a more recent development for journals and publishers covering multiple research disciplines to introduce common requirements for sharing research data, and for reporting the availability of data from their research in published articles [3].

Journal research data policies often include requirements for researchers to provide Data Availability Statements (DAS). The policies of some research funding agencies, such as the UK's Engineering and Physical Sciences Research Council (EPSRC), also require that researchers' publications include DAS. A DAS provides a statement about where data supporting the results reported in a published article can be found, whether those data are available publicly in a data repository, available with the published article as supplementary information, available only privately, upon request or not at all. DAS are often in free-text form, which makes it a non-trivial task to automatically identify the degree of data availability reported in them. This is one of the novel contributions of our study. While DAS can appear in different styles and with different titles depending on the publisher, they are a means to establish and assess compliance with data policies [4–6]. DAS are also known as Data Accessibility Statements, Data Sharing Statements and, in this study, 'Availability of supporting data' and 'Availability of data and materials' statements.

Research data policies of funding agencies and journals can influence researchers' willingness to share research data [7, 8], and strong journal data sharing policies have been associated with increased availability of research data [9]. However, surveys of researchers have also shown that researchers feel they should receive more credit for sharing data [10]. Citations (referencing) in scholarly publications provide evidence for claims and citation counts also remain an important measure of the impact and reuse of research and a means for researchers to receive credit for their work.

Several studies explored compliance with journal data sharing policies [11–15]. For example, DAS in PLOS journals have been found to be significantly on the rise, after a mandated policy has been introduced, even if providing data in a repository remains a sharing method used only in a fraction of articles [16]. This is a known problem more generally: DAS contain links to data (and software) repositories only too rarely [17–19]. Nevertheless there are benefits to data sharing [20–22]. It is known that, for example, the biomedical literature in PubMed has shown clear signs of improvement in the transparency and reproducibility of results over recent years, including sharing data [23]. Some previous studies have shown that, mostly in specific research disciplines—such as gene expression studies [24, 25], paleoceanography [26], astronomy [27] and astrophysics [28]—sharing research data that support scholarly publications, or linking research data to publications, are associated with increased citations to papers [29]. However, to our knowledge, no previous study has sought to determine if providing a DAS, and specifically providing links to supporting data files in a DAS, has an effect on citations across multiple journals, publishers and a wide variety of research disciplines. Making data (and code) available increases the time (and presumably cost) taken to publish papers [30], which has implications for authors, editors and publishers. As more journals and funding agencies require the provision of DAS, further evidence of the benefits of providing them, for example as measured through citations, is needed.

In this study, we consider DAS in journal articles published by two publishers: BMC and PLOS. We focus on the following two questions:

1. are DAS being adopted as per publisher's policies and, if so, can we qualify DAS into categories determined by their contents? In particular, we consider three categories: data

Source and license for this excerpt

The preceding page is page 2 extracted from:
Giovanni Colavizza, Iain Hrynaszkiewicz, Isla Staden,
Kirstie Whitaker, Barbara McGillivray (2020).
The citation advantage of linking publications to research data.
PLOS ONE 15(4): e0230416. https://doi.org/10.1371/journal.pone.0230416
Copyright (c) 2020 Colavizza et al. Licensed under CC BY 4.0:
https://creativecommons.org/licenses/by/4.0/
Changes: the article was reduced to page 2; interactive links were removed;
this attribution page was added. The article text was not edited.
This sample was assembled by Scanfile. Neither the authors nor PLOS
endorse Scanfile or any machine translation of this material.
The article is provided without warranties under its license.