



Current and planned adoption of data sharing policies by editors of Korean scholarly journals

Soo Young Kim¹, Hyun Jung Yi², Sun Huh³

¹Department of Family Medicine, Gangdong Sacred Heart Hospital, College of Medicine, Hallym University, Seoul; ²Medical Library, Hanyang University Guri Hospital, Guri; ³Department of Parasitology and Institute of Medical Education, College of Medicine, Hallym University, Chuncheon, Korea

Abstract

Purpose: This study analyzed the present status of data sharing policies and attitudes towards such policies through a web-based survey of editors of scholarly journals published in Korea.

Methods: From December 26, 2018 to January 3, 2019, a survey was distributed to 1,055 persons listed in the member directories of both the Korean Council of Science Editors and the Korean Federation of Science & Technology Societies. The survey contained four items on subjects' information, three items that gathered information about the journals, and two further items on reasons for adopting or not adopting a data sharing policy and further opinions about such policies.

Results: Of the 100 respondents (from 100 journals), 13 stated that their journals had already adopted a data sharing policy. The strength of the policy was recommendation-only in 10 of those 13 journals. The most frequent reason for adopting a data sharing policy was to follow international trends. The repository sites were the Harvard Dataverse for two journals and Mendeley Data for one. The most common reasons for not adopting a data sharing policy were a lack of knowledge on data sharing, the possibility that submitters would not want to share their data, and the questionable effect of data sharing on scientific development.

Conclusion: Data sharing policies were uncommon among Korean scholarly journals. The advantages and disadvantages of adopting such policies should be discussed more actively among editors and researchers. Furthermore, data sharing infrastructure and training courses are required for data sharing policies to be established in scholarly journals in Korea.

Keywords

Data availability; Journal editor; Knowledge; Republic of Korea

Received: February 4, 2019

Accepted: February 10, 2019

Correspondence to Sun Huh
shuh@hallym.ac.kr

ORCID

Soo Young Kim

<https://orcid.org/0000-0002-3205-9408>

Hyun Jung Yi

<https://orcid.org/0000-0003-0663-4373>

Sun Huh

<https://orcid.org/0000-0002-8559-8640>

Introduction

Data sharing is the practice of making data used for scholarly research available to other investigators. Funding agencies in Korea still do not consider data sharing policies as part of their decision-making process, whereas the United States National Institutes of Health (NIH) recommends data sharing, as follows [1]:

In NIH's view, all data should be considered for data sharing. Data should be made as widely and freely available as possible while safeguarding the privacy of participants, and protecting confidential and proprietary data. To facilitate data sharing, investigators submitting a research application requesting \$500,000 or more of direct costs in any single year to NIH on or after October 1, 2003 are expected to include a plan for sharing final research data for research purposes, or state why data sharing is not possible.

The United States National Science Foundation (NSF) also recommends data sharing policies, as follows [2]:

Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants.

Furthermore, the International Committee of Medical Journal Editors (ICMJE) recommends a data sharing statements follows [3]:

ICMJE will require the following as conditions of consideration for publication of a clinical trial report in our member journals: As of July 1, 2018 manuscripts submitted to ICMJE journals that report the results of clinical trials must contain a data sharing statement as described below. Clinical trials that begin enrolling participants on or after January 1, 2019 must include a data sharing plan in the trial's registration. The ICMJE's policy regarding trial registration is explained at www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html. If the data sharing plan changes after registration this should be reflected in the statement submitted and published with the manuscript, and updated in the registry record.

ICMJE's data sharing statement and data sharing plan is different with general data sharing policy of NIH or NSF, which is not a recommendation of but a statement or plan of data sharing. It is believed that data sharing policies can be beneficial for researchers, as data sharing enables the replication of data analysis and the use of other analytical methodologies for the same data. Biomedical researchers who work on DNA or

protein sequences are already expected to submit their data to National Center for Biotechnology Information (NCBI) databases, including GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>). Therefore, this NCBI database has become invaluable for the study of genes and proteins. Ecological data sharing appeared in the 1950s and became a prerequisite for acquiring funding from a number of funding agencies. As a result, the big data provided by Global Biodiversity Information Facilities or the National Ecologic Observatory Network have become essential to ecological researchers [4].

Besides the data sharing policies of specific research groups or funding agencies, a 2011 report described the adoption of data sharing policies by high-impact factor journals [5]. Of the analyzed journals, 88% had a statement regarding a data sharing policy; however, there was wide variation in the level of the requirements. Since 2016, medical editors of Korean journals that follow ICMJE recommendations have begun to adopt data sharing policies [6]. However, it is unknown how many journals in Korea have adopted a data sharing policy. Furthermore, it is still uncertain whether such policies are a trend in scientific journal publishing beyond the medical field, and no consensus exists among scholarly journal editors throughout the world on the adoption of data sharing policies beyond the clinical data sharing policies found in the medical field. Although data sharing policies remain an unfamiliar concept to some editors of journals in non-medical fields in Korea, it is time for journal editors to consider and discuss such policies.

This study aimed to analyze the present status of adoption of data sharing policies and attitudes towards such policies among editors of scholarly journals published in Korea through a web-based survey. The findings may provide a reasonable and effective basis for introducing data sharing policies to scholarly journals. As a null hypothesis, we hypothesized that there would be no difference in the adoption of data sharing policies across research fields.

Methods

Ethics approval: An anonymous web-based survey was used in this study. It was not possible to identify personal information; furthermore, the survey focused on journal policies. During the survey, no identifiable or sensitive information was collected. Therefore, informed consent was not required according to Korean law [7].

Study design: This study had a cross-sectional and observational design based on a questionnaire survey.

Setting/participants: From December 26, 2018 to January 3, 2019, a survey was distributed to 1,055 persons who were listed in the member directories of both the Korean Council of

Science Editors and the Korean Federation of Science & Technology Societies. The recipients invited to take the survey via email were either journal editors or societies' official email recipients. A Google survey was used. The survey site was <http://bit.ly/2GwuisA>. Survey items included follows: (1) field, (2) role, (3) gender, (4) year of work experience, (5) language of journal, (6) year of launch of journal, (7) status of data sharing policy, (8) reasons for (not) adopting a data sharing policy, (9) plans to adopt a data sharing policy, and (10) further opinions.

Validity and reliability of the acceptability questionnaire: The 4 items that gathered information on subjects' characteristics and the 3 items that collected journal information were basic items that were not subject to bias. We were not able to identify studies in the literature describing the results of surveys of editors about data sharing policies. However, a previous study presented survey results on sharing of clinical trial data from clinical trialists, although the 32 items used in the survey were focused on clinical trialists, not editors [8].

The questionnaire items on editors' reasons for adopting or not adopting a data sharing policy, their plans to adopt such a policy, and further opinions were agreed upon as suitable for the aims of the survey through a discussion among the 3 co-authors of this study. Since the responses to items were not on Likert scales, a reliability test was not done.

Variables: The variables related to the characteristics of respondents and their journals are listed in Table 1. The questionnaire items were considered to be the outcome.

Data sources/measurement: The source of all variables was response data from the survey questionnaire. The measurement method was a comparative analysis.

Bias: There was no noteworthy source of bias during the collection of the survey results or the analysis.

Study size: The sample size ($n = 100$) was 9.5% of all invited subjects. If the effect size was set to 0.7; alpha error probability, 0.05; power, 0.95; and estimated allocation ratio, 1, an adequate sample size for a comparative analysis was estimated as 90 based on GPower ver. 3.1.9.2 [3].

Quantitative variables: All variables were nominal and quantitative. There was no qualitative analysis.

Statistical methods: A descriptive analysis was carried out. The respondents were compared by research field using DB-STAT ver. 5.0 (DBSTAT Co., Chuncheon, Korea), available from <http://dbstat.com>.

Results

All responses from the subjects and the coded content are available in Dataset 1. The characteristics of the 100 respondents and information about their journals are summarized

Table 1. Characteristics of respondents and their journals (from 100 Korean editors)

Item	Option	Number
Fields	Medical/health	54
	Engineering	22
	Natural science	13
	Agricultural science & fisheries	6
	Social science & humanities	3
	Multi-disciplinary	2
Role	Editor	93
	Associate editor	4
	Managing editor	1
	Editorial assistant	2
Sex	Male	71
	Female	29
Experience in current role	Less than 1 yr	8
	1–2 yr	12
	3–6 yr	36
	Over 6 yr	44
Language of journal	English	52
	English or Korean	33
	Korean	15
	Other	0
Year of launch	1945–1949	1
	1950–1959	5
	1960–1969	10
	1970–1979	15
	1980–1989	24
	1990–1999	13
	2000–2009	24
2010–present	8	
Present data sharing policy	Yes	13
	No	87

in Table 1. Medical/health editors comprised 54% of the respondents. The vast majority of the respondents were editors (93%). There were more men (71%) than women. Most respondents (80%) had served in their role for over 2 years, and as many of 44% of editors reported having over 6 years of editorial experience for their journal. Fifty-two percent of the journals had an English-only language policy, and 69% were launched in 1980 or later. Thirteen journals had already adopted a data sharing policy at the time of the survey. As shown in Fig. 1, 76.9% of those journals were in the medical/

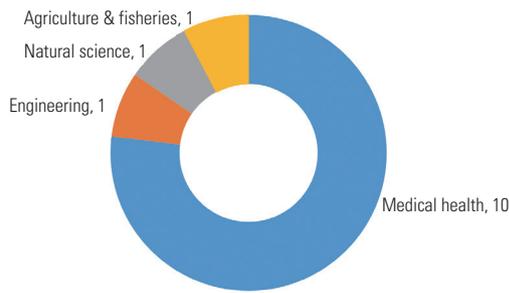


Fig. 1. Fields of the 13 journals in Korea that have adopted a data sharing policy.

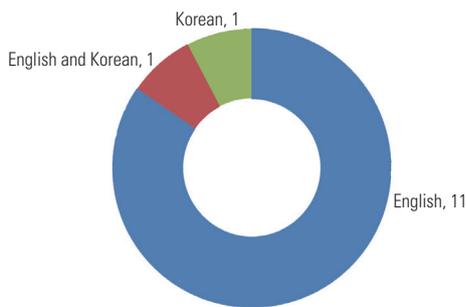


Fig. 2. Language of the 13 journals in Korea that have adopted a data sharing policy.

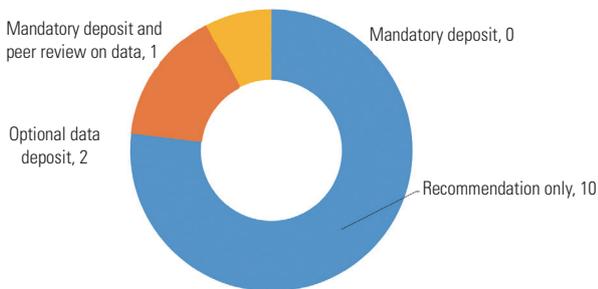


Fig. 3. Strength of the data-sharing policies of the 13 journals in Korea with such a policy.

health field. Of the 13 journals with a data sharing policy, 11 were published only in English (Fig. 2). Although a comparative analysis of the proportion of journals that had adopted a data sharing policy according to field was done, it was difficult to reach any conclusions about whether differences were statistically significant due to the lack of a normal distribution of the data from each field. However, our findings indicate that journals in the medical/health field had adopted data sharing policies more broadly than those in other fields. The strength of the data sharing policies is presented in Fig. 3. The data sharing policies were recommendation-only in the majority of cases (76.9%) in Korea. The reasons for adopting data

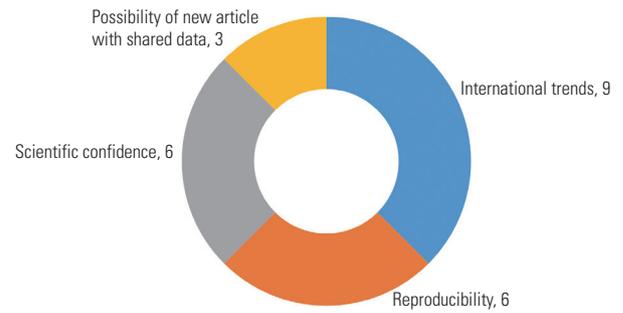


Fig. 4. Reasons for adopting a data sharing policy, among the 13 journals in Korea with such a policy.

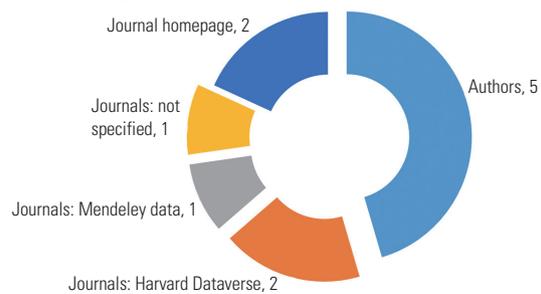


Fig. 5. Who has the responsibility for data deposit and which repository sites are used by the journals in Korea that have adopted a data sharing policy.

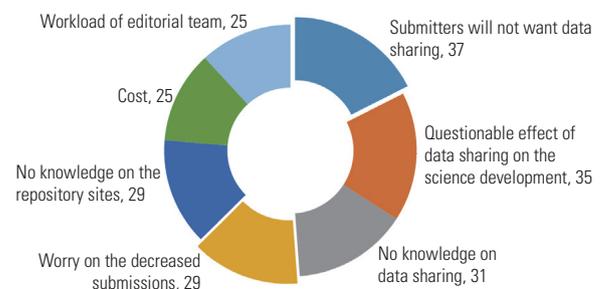


Fig. 6. Reasons for not adopting a data sharing policy, among the 87 journals in Korea with no such policy.

sharing policies are shown in Fig. 4; multiple choices were possible, but the main reasons were reproducibility (6) and scientific confidence (6). Other reasons included increased brand value and protection against falsification or fabrication. The repository sites were selected by the authors (5) or journals (4), and in 2 cases, the data were deposited on the journal's homepage (2). The repository sites chosen by the journals were the Harvard Dataverse (2) and Mendeley Data (1); in the remaining case, the repository site was not clear (Fig. 5).

The reasons for not adopting a data sharing policy are listed in Fig. 6. The following reasons were the most frequent responses: submitters will not want to share their data (37); the

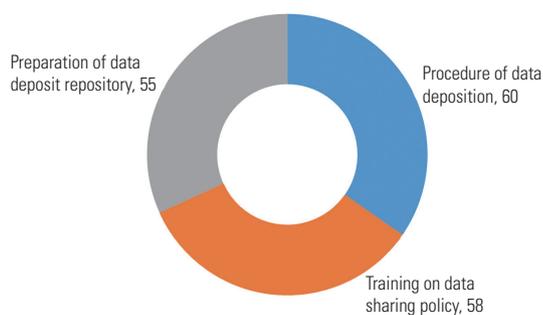


Fig. 7. Factors identified as necessary for editors to establish data sharing policies in Korea.

questionable effect of data sharing on scientific development (35); and a lack of knowledge on data sharing (31). Some of other reasons not listed in Fig. 6 included copyright problems and the lack of critical consideration by the society. Of the 87 editors of journals that had not adopted a data sharing policy, 36 had a plan to adopt such a policy, while 49 said that they had no plan to do so. Two editors did not respond to this item. Fig. 7 shows the editors' responses regarding factors that would be necessary for data sharing policies to be established (multiple choices were possible). Preparation of a data deposit repository (55), data deposit procedures (60), and training on data sharing policies (58) were selected by 87 editors. Other desires included budget preparation and a process of agreement within the society.

Discussion

Key results: Thirteen of the 100 respondents reported that their journal had adopted a data sharing policy (Figs. 1, 2). The data sharing policies were recommendation-only in most cases (76.9%) (Fig. 3). Three journals mentioned a data repository: two Harvard Dataverse and one Mendeley Data (Fig. 5). Data sharing policies are still unfamiliar to some Korean editors (31%) (Fig. 6). Of the 87 editors whose journals had not adopted a data sharing policy, 41.2% had a plan to adopt such a policy.

Interpretation and suggestions: We found daunting circumstances regarding data sharing policies in scholarly journals in Korea. In contrast, numerous well-known foreign journals have adopted data sharing policies. For example, a previous study assessed the presence of data sharing and code sharing policies in 170 journals in the categories of mathematical and computational biology, statistics and probability, and multidisciplinary sciences in Web of Science, with the following results: "of 170 journals, 38% had a data policy, 22% had a code policy, and 66% had a supplemental materials policy as of June 2012" [9]. Editors also should decide the strength of the

policy. If it is only a recommendation, it is difficult to recruit enough submissions with shared data. Among the respondents, only one journal editor stated that the shared data were reviewed for studies at the present time. Furthermore, data files must be reviewed in light of the following considerations: who submits the data; for how long the data will be shared; where the data will be deposited; what portion of the data will be deposited; and what format of data will be deposited. No consensus exists regarding data format in a variety of fields, except for gene and protein data, which should be deposited at the NCBI. Journal editors should jointly discuss a standardized format. Some editors mentioned the possibility of new articles based on shared data (Fig. 4). Work to clarify the merits of this possibility should be encouraged. For example, if analyses of shared data are included more often in meta-analyses or systematic reviews, that would potentially provide evidence for the merits of data sharing policies. Only three editors deposited data to repository sites such as the Harvard Dataverse or Mendeley Data (Fig. 5). According to Korean law, limitations exist on transferring sensitive information of human subjects to data repositories in foreign countries; such information includes "ideologies, beliefs, admission to or withdrawal from a trade union or political party, political opinions, health, sexual life, and other personal information that is likely to threaten the privacy of any data subject noticeably" based on the Enforcement Decree of the Personal Information Protection Act [10]. Although legal debates continue regarding the definition of sensitive information, it would be preferable to provide a public data repository site in Korea to avoid these legal issues.

Lack of knowledge about data sharing policies and data repository were frequently reported reasons for not adopting such policies; the frequency of these responses suggests that some editors were not aware of data sharing policies (Fig. 6). It is time to inform them about such policies through training courses or workshops organized by editors' organizations, such as the Korean Council of Science Editors. Before deciding whether or not to adopt a data sharing policy, the advantages and disadvantages of doing so should be discussed actively among editors and researchers at editors' workshops. Some funding agencies, including the US NIH, recommend a data sharing policy, although it is not currently mandatory. Furthermore, the ICMJE recommendation regarding clinical data sharing is not mandatory. It requires a statement by authors indicating whether the data may be shared. Therefore, if authors do not want to share their data, it is acceptable to state that decision in the data sharing statement. However, if authors want to share their data, editors and publishers should consider an appropriate data repository and ensure data review. It may be challenging to introduce such policies if tech-

nical difficulties emerge during adoption. To remove technical barriers, editors and publishers should also receive appropriate training.

Limitations: The data were not collected from randomly selected samples. There may be non-response bias. Therefore, it is difficult to state that the results of this study represent the current situation in all science journals in Korea. Second, only 13 respondents stated that their journal had a data sharing policy. Therefore, it was challenging to conduct a comparative analysis according to characteristics of the respondents or their journals due to the non-normal distribution of the data. The null hypothesis that there would be no difference in the adoption of data sharing policies across research fields could not be rejected due to the non-normal distribution of the sample.

Generalizability: Since the sample was not randomized, care should be taken when extrapolating our findings to represent all scholarly journals in Korea. There are 643 scientific journals in Korea according to the Korea Citation Index, available from: https://www.kci.go.kr/kciportal/po/statistics/poStatisticsMain.kci?tab_code=Tab1 (cited 2019 Feb 4). More intensive data collection is necessary to characterize the present situation. The data described the present situation and trends in the adoption of data sharing policies by journals in the near future. According to our results, the possibility of such policies being adopted is promising.

Conclusion: Publishing societies and organizations in Korea should decide whether to adopt a data sharing policy. According to our results, only 13% of journals had adopted such a policy and 49% of editors did not have a plan to do so. Before making such decisions, training courses on data sharing are required in order to help editors understand such policies more lucidly. Furthermore, infrastructure (such as establishment of a domestic data repository) is also required to support editors who would like to adopt such a policy.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Data Availability

Dataset 1. Response data of the questionnaire survey and the content of coding is available from the Harvard Dataverse at: <https://doi.org/10.7910/DVN/F41EQP>.

References

1. National Institute of Health. NIH data sharing policy and implementation guidance [Internet]. Bethesda, MD: National Institute of Health; c2003 [cited 2019 Feb 1]. Available from: https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm
2. National Science Foundation. Dissemination and sharing of research results [Internet]. Alexandria, VA: National Science Foundation; c2011 [cited 2019 Feb 1]. Available from: <https://www.nsf.gov/bfa/dias/policy/dmp.jsp>
3. Taichman DB, Sahni P, Pinborg A, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. *Ethiop J Health Sci* 2017;27:315-8. <https://doi.org/10.4314/ejhs.v27i4.2>
4. Michener WK. Ecological data sharing. *Ecol Inform* 2015; 29:33-44. <https://doi.org/10.1016/j.ecoinf.2015.06.010>
5. Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JP. Public availability of published research data in high-impact journals. *PLoS One* 2011;6:e24357. <https://doi.org/10.1371/journal.pone.0024357>
6. Huh S. Establishment of an open data policy for Journal of Educational Evaluation for Health Professions, appreciation for invited reviewers, and acknowledgement of volunteers who made audio recordings. *J Educ Eval Health Prof* 2017;14:37. <https://doi.org/10.3352/jeehp.2017.14.37>
7. Ministry of Government Legislation. Enforcement Rule of Bioethics and Safety Act [Internet]. Sejong: National Law Information Center; 2017 [cited 2019 Jan 19]. Available from: <http://www.law.go.kr/lumLsLinkPop.do?lspttninfSeq=75929&chrClsCd=010202>
8. Rathi V, Dzara K, Gross CP, et al. Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ* 2012; 345:e7570. <https://doi.org/10.1136/bmj.e7570>
9. Stodden V, Guo P, Ma Z. Toward reproducible computational research: an empirical analysis of data and code policy adoption by journals. *PLoS One* 2013;8:e67111. <https://doi.org/10.1371/journal.pone.0067111>
10. Ministry of Government Legislation. Enforcement Decree of the Personal Information Protection Act [Internet]. Sejong: Korea Legislation Research Institute; 2017 [cited 2019 Jan 19]. Available from: http://elaw.klri.re.kr/kor_service/lawView.do?lang=ENG&hseq=45683