

Comparison of length limits and the actual length of abstracts in pharmacology, oncology, and neurology journals listed in PubMed

Eungi Kim, Yong-Gu Lee

Department of Library and Information Science, Keimyung University, Daegu, Korea

Abstract

Purpose: This study aimed to compare the length limits specified in the author guidelines with the actual length of abstracts in 90 journals in the fields of pharmacology, oncology, and neurology. Specifically, the following parameters were examined: abstract formats among the three subject areas; the relationship between the length limit and the actual length of abstracts; and actual abstract length according to the number of subheadings, the length of structured abstract subheadings, the length of frequently used subheading sets, and clinical trial registration information.

Methods: Thirty journals from each of three medical fields (pharmacology, oncology, and neurology) were selected from Elsevier's Scimago Journal Rank. This included the journals indexed in PubMed from 2018 to 2019 that published the most articles. Article abstracts from these journals were used to create a dataset for this study. Descriptive, comparative, and correlational analyses of data for the three fields were conducted.

Results: The number of subheadings and abstract length increased in parallel. The *Results* component was the longest, suggesting that authors tended to use longer text to report results than for other structural abstract components. Authors generally utilized the length limit to a full extent without exceeding it.

Conclusion: The traditionally used 250-word length limit should be reconsidered for pharmacology, oncology, and neurology journals because it disregards the distinctive characteristics of abstracts and length differences between structured and unstructured abstracts. Various characteristics of abstract lengths presented in this study should be considered to establish more justifiable policies.

Keywords

Bibliometrics; Publications; PubMed

Received: November 17, 2020
Accepted: January 22, 2021

Correspondence to Yong-Gu Lee
yonggulee@kmu.ac.kr

ORCID
Eungi Kim
<https://orcid.org/0000-0001-7314-1726>
Yong-Gu Lee
<https://orcid.org/0000-0002-0192-8711>

Introduction

Background/rationale: Along with adhering to other guidelines, an abstract should not exceed the number of words specified by the editorial policy of a journal. Abstracts can be broadly classified into two major formats: unstructured and structured. The major difference between the two formats is that unstructured abstracts (UAs) do not use subheadings and are typically written in a single-paragraph format, while structured abstracts (SAs) use pre-defined subheadings. SAs are more commonly used in medical research [1] and various subheadings are used for SAs based on the introduction, methods, results, and discussion (IMRaD) structure [2]. For example, *Background* is more commonly used than *Introduction*. It is generally believed that an appropriate abstract limit is 250 words or fewer [3]. Although several empirical studies have revealed that SAs are longer than UAs on average [4], abstract length has not been examined thoroughly regarding adherence to the length limit. Linder [5] pointed out that length limits vary substantially across disciplines. Silverberg and Ray [6] showed that most published abstracts exceeded the established length limit in the top five medical journals.

In this study, we examined 90 pharmacology, oncology, and neurology journals indexed in PubMed from 2018 to 2019, focusing on the journals that published the most articles, since these three areas had the highest output in the medical field and thus could provide a comprehensive and representative sample of medical journal abstracts, the analysis of which would yield insights applicable to medical abstracts in general. **Objectives:** Based on abstracts from pharmacology, oncology, and neurology journals, we aimed to examine the following parameters: abstract formats among the three subject areas; the relationship between the length limit and the actual length of abstracts; and actual abstract length according to the number of subheadings, the length of the SA subheadings, length of the frequently used subheading sets, and clinical trial registration information.

Methods

Ethics statement: This study was conducted through a database-based, bibliometric search and analysis of the literature. Neither institutional review board approval nor informed consent was required.

Study design: This quantitative descriptive analysis was conducted to compare the length limits specified in journals' author guidelines with the actual length of abstracts.

Data sources/measurement: We selected three distinctive medical subject areas (pharmacology, oncology, and neurology) from Elsevier's Scimago Journal Rank. These subject areas

shared the common characteristic of having more than 250 listed journals. To consistently ensure a sufficient number of articles across all journals, the top 30 journals indexed in PubMed (i.e., those that published the most articles in each subject area) were selected. The details of the dataset and all 90 journals (30 journals for each of the three subject areas) used in this study are shown in Dataset 1 and 2. Journal article abstracts from 2018 to 2019 were downloaded from PubMed. An equal number of abstracts from journals in each subject area was obtained by identifying the journal with the smallest number of abstracts. To create a dataset for this study, we selected the same number of abstracts from the rest of the journals. For example, *Brain Pathology* had the lowest number (140) in neurology; thus, 140 abstracts were randomly selected from the rest of the journals in neurology. When performing frequency counts of subheadings, words in the singular form were changed to the plural (e.g., "Method" to "Methods").

Statistical methods: The descriptive, comparative, and correlational analysis was done with Excel (Microsoft, Redmond, WA, USA).

Results

Comparison of abstract formats among the three subject areas: Table 1 shows the frequency count of abstract formats based on the editorial policy. A considerable number of journals could be described as mixed, since many journals did not require exclusively one form of abstracts. In the dataset, more UAs than SAs across all subject areas were present; in total, only 26 (28.9%) journals exclusively used UAs and 10 (11.1%) journals exclusively used SAs. Thus, most journals were mixed—that is, they contained both UAs and SAs ($n = 54$, 60%)—although most editorial policies specified the usage of exclusively one type of abstract. Thirty-four (37.8%) journals had editorial policies specifying the use of UAs, while 43 (47.8%) journals had policies requiring SAs. Only two (2.2%) journals used semi-SAs. Eight journals (8.9%) did not specify the abstract format, despite having statements on the abstract length limit. For the journals that specified the required range of abstract length, rather than the maximum length, the upper bound was used as the length limit (e.g., a 200-word length limit for a required range of 100 to 200 words).

Comparison between the length limit and actual abstract length: The length limit was compared with the actual length of abstracts based on the journals that exclusively used SAs or UAs. Excluding the mixed abstracts allowed us to examine authors' tendencies regarding individual journals' abstract limits. Fig. 1 shows the lengths of UAs and SAs, excluding subheadings, for the journals that had statements on the ab-

Table 1. Comparison of abstract formats among the three subject areas

Criteria	Abstract format	Subject area			
		Oncology (n=30)	Neurology (n=30)	Pharmacology (n=30)	Total (n=90)
Editorial policy (abstract type)	Unstructured	14 (46.7)	13 (43.3)	7 (23.3)	34 (37.8)
	Structured	14 (46.7)	13 (43.3)	16 (53.3)	43 (47.8)
	Semi-structured ^{a)}	0 (0.0)	1 (3.3)	1 (3.3)	2 (2.2)
	Not specified	2 (6.7)	2 (6.7)	4 (13.3)	8 (8.9)
	Mixed ^{b)}	0 (0.0)	1 (3.3)	2 (6.7)	3 (3.3)
	Total	30 (100)	30 (100)	30 (100)	90 (100)
Dataset (actual length)	Unstructured	12 (40.0)	8 (26.7)	6 (20.0)	26 (28.9)
	Structured	2 (6.7)	5 (16.7)	3 (10.0)	10 (11.1)
	Mixed ^{c)}	16 (53.3)	17 (56.7)	21 (70.0)	54 (60.0)
	Total	30 (100)	30 (100)	30 (100)	90 (100)

Values are presented as number (%).

^{a)}Semi-structured abstracts have a single-paragraph format, with each sentence corresponding to a section in the paper. When calculating the abstract length, semi-structured abstracts were counted as unstructured abstracts due to missing subheadings; ^{b)}Journals allowing both unstructured and structured abstracts; ^{c)}Journals containing both unstructured and structured abstracts.

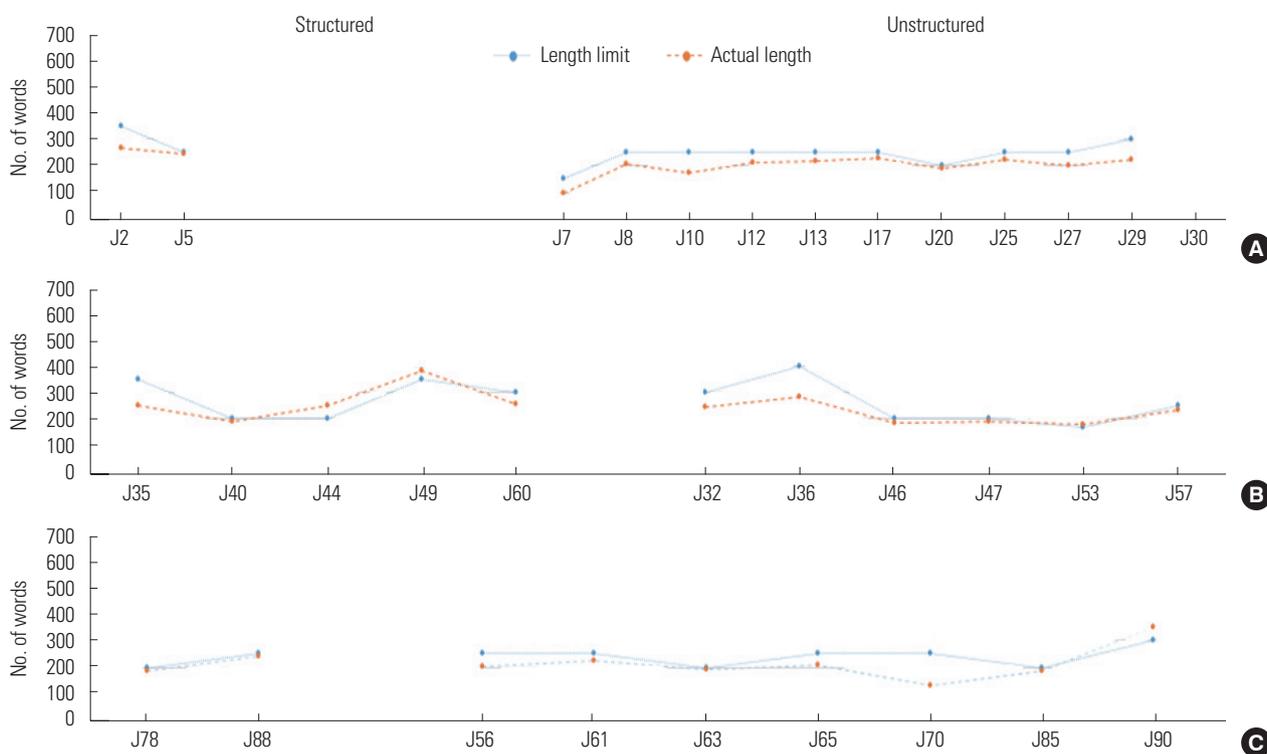


Fig. 1. The abstract lengths of selected journals. The names of the journal index numbers are given in Dataset 1. (A) Oncology, (B) neurology, and (C) pharmacology.

abstract length limit and exclusively used either UAs or SAs. The journals shown on the left side of Fig. 1 exclusively used SAs (more than 99%), while those on the right side exclusively used UAs (more than 99%). *The Cerebellum* (J38) was exclud-

ed from this figure, as the journal required SAs, but mostly consisted of UAs in the dataset. As shown, the range of length limits varied widely (150 to 400 words). Slightly below the length limits, a wide range of actual lengths (98 to 387 words)

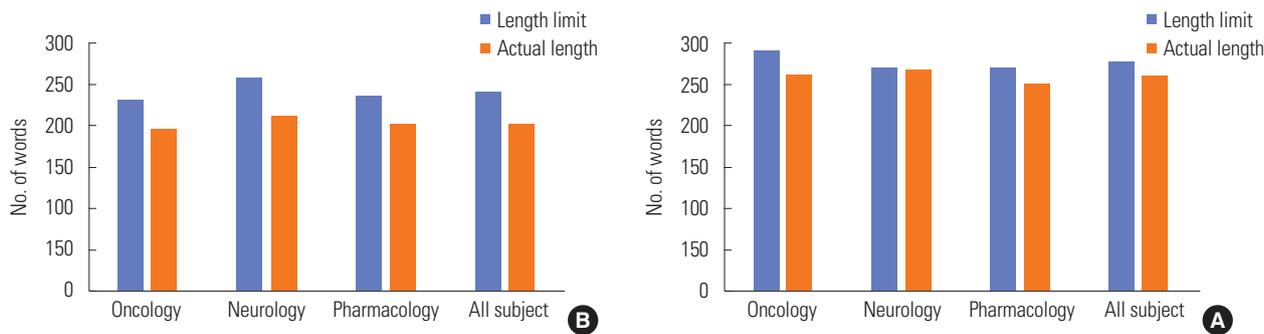


Fig. 2. Overall comparison between length limit and actual length. (A) Structured and (B) unstructured abstracts.

Table 2. Number of subheadings and actual abstract length

No. of subheadings	No. of unstructured abstracts (average word count)				No. of structured abstracts (average word count)						
	0	1	2	3	4	5	6	7	8	9	
Subject area											
Oncology	4,532 (195.8) ^{a)}	578 (233.5) ^{a)}	2 (210)	114 (233.2) ^{a)}	3,162 (259.9) ^{a)}	368 (310.8) ^{a)}	2 (276)	2 (280)	Null	Null	
Neurology	2,167 (211.9) ^{a)}	144 (255) ^{a)}	7 (297.7)	201 (196.9) ^{a)}	1,114 (258.8) ^{a)}	426 (281.9) ^{a)}	73 (372.6)	42 (397.3)	26 (398.1)	Null	
Pharmacology	2,775 (202.6) ^{a)}	37 (234.5)	12 (240.3)	321 (218.8) ^{a)}	2,498 (249.8) ^{a)}	542 (262.3) ^{a)}	90 (301.1)	228 (716.1) ^{a)}	3 (365.7)	4 (630.8)	

Null indicates that no case was found for the particular cell.

^{a)}Instances with sufficiently large numbers of abstracts (100 or more).

could be also noticed. By and large, SAs were longer than UAs. The actual length exceeded the length limit in some journals (e.g., J44, J49, J50).

The overall line patterns show that the actual length fluctuated slightly along with the length limit. Furthermore, we found positive correlations between the length limit and actual length in all subject areas: oncology ($r=0.83$), neurology ($r=0.76$), and pharmacology ($r=0.67$).

In addition, the length limit was compared with the actual length regardless of the abstract format. Unlike Fig. 1, where the result was produced based on only selected journals, Fig. 2 was produced using all 90 journals (i.e., 30 journals in each subject area). This figure indicates that the actual lengths were mostly below the length limits in all subject areas. Furthermore, SAs were longer than UAs in all subject areas, particularly in terms of actual length as opposed to length limits. SAs in neurology showed the smallest difference between the length limit and actual length (3.5 words), whereas UAs in neurology showed the greatest difference between the length limit and actual length (45.8 words). The difference between the length limit and actual length for all subject areas was greater in UAs than in SAs. For UAs, the average length limit (242.4 words) was greater than the actual length (203.5 words) by 19.1%, while for SAs, the overall average length

limit (275.9 words) was greater than the actual length (258.7 words) by only 6.6%.

Actual abstract length according to the number of subheadings:

Journals used a varying number of subheadings. Table 2 shows that the number of SA subheadings in abstracts from the sampled journals ranged from one to nine. SAs having less than four or more than five subheadings were less common, and most SAs used three to five subheadings. In oncology, a substantial number of abstracts contained only one subheading, such as *Significance*. The length generally increased with the number of subheadings. However, this occurred only when a sufficient number of abstracts was used to obtain the abstract length. For all three medical subject areas, the length increased steadily as the number of subheadings increased from three to five. Within this number of subheadings, oncology journals showed the highest number of abstracts with the greatest length: 233 words for three subheadings and 310 words for five subheadings. The overall average length was 216.3 words for three subheadings, 256.2 words for four subheadings, and 285 words for five subheadings. On average, the length increased by 39.9 words (18.0%) when the number of subheadings increased from three to four. The length increased by 28.8 words (11.3%) when the number of subheadings increased from four to five. The length of UAs is

Table 3. Components of structured abstracts and their length

Rank	Frequency	Subheading	Oncology	Neurology	Pharmacology	All subject areas
1	8,020	Results	105.5±35.9	101.2±43.2	95.4±35.3	101.0±37.5
2	7,754	Conclusions	38.9±15.9	41.3±17.9	40.5±17.1	39.9±16.7
3	7,113	Methods	67.1±26.5	71.4±36.3	66.3±27.7	67.7±29.4
4	4,571	Background	52.9±22.7	46.5±25.2	56.4±29.9	53.2±25.8
5	2,319	Objectives	46.5±21.8	36.7±22.3	35.0±18.5	37.9±21.3
6	1,555	Purpose	48.7±21.1	26.0±4.20	49.2±23.9	49.0±22.9
7	573	Introduction	52.9±22.4	64.5±32.8	53.9±24.4	53.9±24.2
8	509	Materials and methods	60.4±22.6	69.3±36.0	56.0±24.2	60.2±22.8
9	453	Implications	30.7±10.2	Null	47.2±20.1	37.3±17.0
10	352	Significance	27.8±7.30	49.2±17.5	19.0±UNDEF	33.5±14.5

Values are presented as mean±standard deviation; Null indicates that no case was found for the particular cell.

also displayed in Table 2. As shown, UAs were overall shorter than SAs.

Lengths of SA subheadings: As expected, there was considerable variability in the occurrence and length of SA components. Table 3 shows the top 10 most used SA components and their lengths. Plotting the frequency count of subheadings would suggest an exponential distribution. *Results* (ranked 1st) was used approximately 22 times more than *Significance* (ranked 10th). On average, *Results* was the most frequently used subheading among the subject areas, and it was the longest (101.0±37.5 [standard deviation]). In contrast, *Significance* was the shortest (33.5±14.5 [standard deviation]), although it did not appear in all subject areas. This suggests that authors tended to use a higher number of words to describe their findings than for other components.

Lengths of the frequently used subheading sets: Table 4 shows the details of the top 10 most frequently used subheading sets. The information on subheading sets is divided into two groups: journal abstracts and editorial policy requirements. As shown in Table 4, the most frequently used subheading set was *Background, Methods, Results, Conclusions* (27.4%). In terms of editorial policy, this structure was used in eight out of 90 journals (8.9%). In terms of abstract length, the length limit of abstracts using the *Background, Methods, Results, Conclusions* subheading set was higher (300.0 words) than the actual length (265.7 words) by 12.9%. The *Purpose, Methods, Results, Conclusions* subheading set was also widely used in both groups. If normalization was applied by combining *Objectives* with *Purpose*, the combined subheading set would rank the highest in both groups. The average length limit of SAs in terms of subheading sets was higher than the overall average length limit of SAs.

Abstract length and clinical trial registration information:

Papers involving clinical trials are often required to disclose clinical trial registration information. This study's results showed that most journals required authors to specify the trial registration number or include the phrase "retrospectively registered" in the abstract, although the inclusion format varied among journals. Most journals instructed authors to include statements concerning the clinical information at the end of the abstract, either as a separate subheading ("trial registration") or without the subheading. The required statements were relatively short compared to other information. Out of 18,557 abstracts, 225 UAs (1.2%) and 668 SAs (3.6%) had clinical trial information.

Fig. 3 portrays the relationship between trial registration information and abstract length. This figure shows that, regardless of whether clinical trial information was present, the actual length was less than the average length limit. This inclusion of clinical trial information affected the actual length of SAs more than that of UAs since they contained this information more frequently. SAs containing trial registration information had 39.4 words (17.6%) more than those without this information. UAs followed the same pattern, with abstracts containing trial registration information having 56.3 words (28.0%) more.

Discussion

Interpretation: We empirically compared abstract limits with abstract lengths, taking into account the number of elements, specifically required items, and standard forms (UAs and SAs).

First, a considerable number of journals contained both UAs and SAs despite specific requirements for the form of abstracts in the author guidelines. This indicates that some jour-

Table 4. Top ten frequently used subheading formations and their abstract lengths

Criteria	Rank	Subheading set	No. of papers (%)	No. of journals appearing at least once	Average no. of words
Based on journal abstracts	1	Background, Methods, Results, Conclusions ^{a)}	2,741 (27.4)	45	265.7
	2	Purpose, Methods, Results, Conclusions ^{a)}	960 (9.6)	21	239.9
	3	Objectives, Methods, Results, Conclusions ^{a)}	782 (7.8)	38	259.2
	4	Background, Objectives, Methods, Results, Conclusions ^{a)}	398 (4)	23	235.3
	5	Implications	270 (2.7)	1	229.6
	6	Background and purpose, Methods, Results, Conclusions	261 (2.6)	5	258.2
	7	Objectives, Materials and methods, Results, Conclusions ^{a)}	257 (2.6)	7	279.0
	8	Significance	256 (2.6)	2	222.1
	9	Purpose, Experimental design, Results, Conclusions	220 (2.2)	6	246.4
	10	Objectives, Methods, Results, Interpretation ^{a)}	215 (2.2)	4	776.0
		Total	9,996 (100)	Average	269.8
Based on editorial policy	1	Purpose, Methods, Results, Conclusions ^{a)}	8 (12.5)	-	243.8
	1 (tie)	Background, Methods, Results, Conclusions ^{a)}	8 (12.5)	-	300.0
	3	Objectives, Methods, Results, Conclusions ^{a)}	5 (7.8)	-	275.0
	4	Objectives, Materials and methods, Results, Conclusions ^{a)}	4 (6.3)	-	275.0
	5	Background, Objectives, Methods, Results, Conclusions ^{a)}	3 (4.7)	-	250.0
	6	Purpose, Materials and methods, Results, Conclusions	2 (3.1)	-	250.0
	6 (tie)	Objectives, Methods, Results, Interpretation ^{a)}	2 (3.1)	-	250.0
	6 (tie)	Introduction, Methods, Results, Conclusions	2 (3.1)	-	275.0
	9	Study design, Results, Conclusions	1 (1.6)	-	200.0
	9 (tie)	Statement of problem, Purpose, Materials and methods, Results, Conclusions	1 (1.6)	-	250.0
		Total	64 (100)	Average	296.2

^{a)}The subheading sets that appear in both groups.

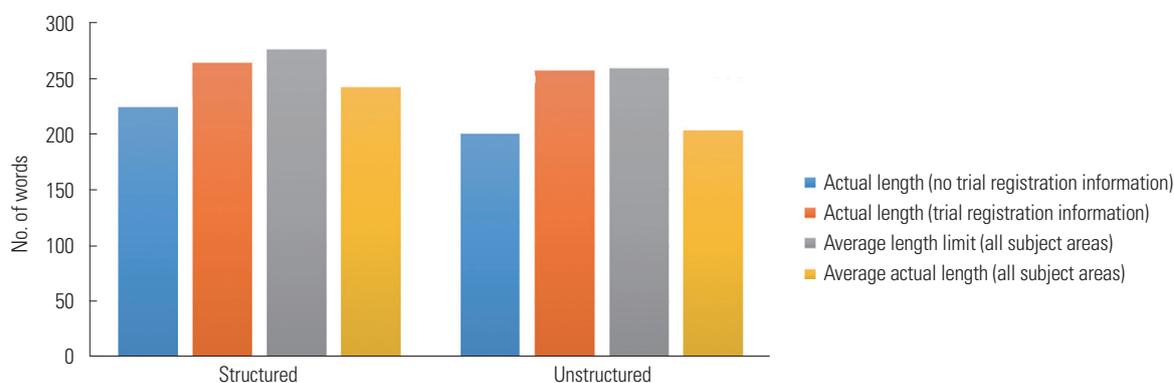


Fig. 3. Clinical trial information and the abstract length of structured and unstructured abstracts.

nals may have loosely enforced their editorial policies regarding the abstract format. Moreover, editorial policies regarding the abstract format may have changed since the publication period of 2018. For instance, *The Cerebellum* required SAs for

new manuscript submissions, but UAs were mostly present in the dataset.

Second, using more subheadings increased the overall lengths of abstracts. The most commonly used numbers of subhead-

ings was four and five, and abstract length increased steadily as the number of subheadings increased from three to five. Similarly, a wide range of subheadings was found by a study conducted by Eid et al. [7].

Third, the longest SA component was *Results*. This suggests that authors generally tended to use longer text to report results. The *Results* subheading also the most frequently appearing component in language and literature journals that used the IMRaD structure [8].

Fourth, positive correlations between the abstract limit and the actual length, along with the differences between required and actual lengths, suggest that authors generally utilized the length limit to a full extent without exceeding it.

Fifth, requiring clinical trial information in the abstract affected the abstract length of UAs more than that of SAs, but the overall length increase was minimal due to the limited numbers of abstracts that had clinical trial information.

Lastly, authors generally followed the journal's editorial policies regarding abstract length by keeping the length slightly below the specified length limit. Overall, the authors utilized the maximum allowed length for SAs. At the same time, there was leeway in terms of accepting journal abstracts that exceeded the specified abstract length. The overall abstract lengths obtained in this study were longer than those reported in previous studies. Atanassova et al. [9] showed an average abstract length of 185.1 words based on seven PLOS journals. PLOS journals presumably use SAs, although the authors did not explicitly state the abstract form. The present study's results showed 203.5 words for UAs and 228.7 for SAs. In contrast to the findings of Silverberg and Ray [6], most abstracts kept their length below the allowable length limit.

Limitation: The actual length and the length limits were taken from slightly different periods. The actual length was based on research articles published from 2018 to 2019, whereas the editorial policies were more up-to-date since they were collected at the time of conducting this study (September 2020). Furthermore, this study was limited to three subject areas and journals selected from each subject area. The results of this study should be interpreted in the context of these limitations.

Conclusion: The traditionally used 250-word length limit should be reconsidered for pharmacology, oncology, and neurology journals because it disregards the distinctive characteristics of abstracts and length differences between SAs and UAs. For new journals in particular subject areas, a pragmatic, yet more systematic approach for establishing a length limit would be to use a certain tolerance level over the empirically obtained average abstract length in the subject area. Furthermore, various characteristics of abstract lengths presented in this study should be taken into account in establishing more justifiable policies.

Conflict of Interest

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

Funding

The authors received no financial support for this article.

Data Availability

Dataset file is available from: the Harvard Dataverse at: <https://doi.org/10.7910/DVN/JKY2YI>

Dataset 1. Journal titles with percentages of structured and unstructured abstracts

Dataset 2. Journal titles and the PubMed ID of the abstracts downloaded

References

1. Pogue B. Structured abstracts: the time has come for the Journal of Biomedical Optics! J Biomed Opt 2019;24:1-2. <https://doi.org/10.1117/1.JBO.24.10.100101>
2. Bahadoran Z, Mirmiran P, Kashfi K, Ghasemi A. The principles of biomedical scientific writing: abstract and keywords. Int J Endocrinol Metab 2020;18:e100159. <https://doi.org/10.5812/ijem.100159>
3. Sollaci LB, Pereira MG. The introduction, methods, results, and discussion (IMRAD) structure: a fifty-year survey. J Med Libr Assoc 2004;92:364-7.
4. Kim E. The lengths of LIS journal abstracts: a comparison of structured abstracts and traditional abstracts. J LISSAS-PAC Jpn 2018;1:46-9.
5. Linder D. English abstracts in electronic translation studies journals in Spain. Paper presented at: TEEM '13: the First International Conference on Technological Ecosystem for Enhancing Multiculturality; 2013 Nov 14-15; Salamanca, Spain. p. 391-8. <https://doi.org/10.1145/2536536.2536596>
6. Silverberg O, Ray JG. Variations in instructed vs. published word counts in top five medical journals. J Gen Intern Med 2018;33:16-8. <https://doi.org/10.1007/s11606-017-4196-6>
7. Eid T, vanSonnenberg E, Azar A, Mistry P, Eid K, Kang P. Analysis of the variability of abstract structures in medical journals. J Gen Intern Med 2018;33:1013-4. <https://doi.org/10.1007/s11606-018-4428-4>
8. Jamar N, Sauperl A, Bawden D. The components of abstracts: the logical structure of abstracts in the areas of materials science and technology and of library and information science. New Libr World 2014;115:15-33. <https://doi.org/10.1108/NLW-01-2014-0011>

doi.org/10.1108/NLW-09-2013-0069
9. Atanassova I, Bertin M, Lariviere V. On the composition

of scientific abstracts. J Doc 2016;72:636-47. <https://doi.org/10.1108/JDOC-09-2015-0111>