

Original Research- Sample work

**An empirical examination of the text structure of
original research articles in medical journals**

Abstract

The structure of the main text, the so-called introduction, methodology, findings, and discussion structure all have a role in the success of an article's publication (IMRAD). The primary goal of the research is to present new findings on the number of paragraphs (pars.) per section utilized in major medical journals articles. The second goal is to study additional structural components, such as the quantity of tables, figures, and references and the accessibility of supplemental information. The data were evaluated from original studies published in The BMJ, The Journal of the American Medical Association, The New England Journal of Medicine, and PLOS Medicine in 2005, 2010, and 2015. 30 articles were examined per journal and year. To get pooled estimates, random effect meta-analyses were used. Linear mixed representations were used to examine the influence of time. The IMRAD framework was followed in all of the articles. For all journals, the number of pars. Per section grew with time, increasing by 1.08 (95 % confidence interval (CI): 0.70–1.46) pars. Every two years. The techniques section had the most growth (0.29 pars. per year; 95 % confidence interval (CI): 0.19–0.39). The journal PLOS Medicine has the most pars. Although the number of tables did not change, the number of figures and references did. To increase the chances of an article presence being published, the standard IMRAD format and the basic layout of the target journal should be employed. Supplementary material is now commonplace. Authors should utilize 3/10/9/8 pars. for the introduction/methods/results/discussion sections if no journal-specific material is provided.

Introduction

One says that scientists frequently encounter "publish and gain." Working in the field of research necessitates a steady stream of publications. Scientists are fiercely competitive, and journal space is limited. However, the world of publishing can be a minefield, and many people find writing difficult (Hochberg & Hochberg, 2019). At the same time, scientific communication is rarely taught, and scientific writing is fundamentally different from literary writing. Only a few authors concentrate on the writing process. Albert developed one such procedural method (Albert, 2018), and he explained article writing using a 10-step process (Albert, 1996). Albert's proposal considers a sales model in one of these phases. The target audience for this sales strategy to publishing is the journal editor because the editor is the gateway for article acceptance or rejection and early rejections known as desk-rejections. The initial impression of a document, as always, is the best one (Fathi M Sherif, 2013). Thrower gave a reason for accepting (Thrower, 2013) and rejecting (Thrower, 2013) papers for publication. The way the content is presented is an essential consideration (Korner, 2008). It is self-evident that writers adhere to the target journal's author guidelines. The reference style also matters, and "a neatly structured work makes the editor pleased since he doesn't have to do anything further from his side in terms of presentation" (Fathi M Sherif, 2013). Froese and Bader (Froese & Bader, 2019) wrote an editorial on the reasons for early article rejections, and they noted that manuscripts that do not fit the format of a typical article are likely to be rejected. In this context, we must consider an article's standard structure.

The introduction, methods, results, and discussion (IMRAD) format is a standard structure for the main text of a medical journal article. Although it was proposed as a typical

structure at the turn of the 20th century, it was not until 1965 that it became the primary structure, and it was the only article structure in the 1980s (Sollaci & Pereira, 2004). The massive change in the mid-1960s was most likely prompted by a conference of medical editors held at the World Medical Association's 19th General Assembly (Brain, 1965), during which Hill (1965) proposed that research articles answer four questions: why did you start, what did you do, what answer did you get, and what does it all mean? There has been little study done on the specifics of these four portions of an article (Albert, 2016; Araujo, 2014). Albert (2016) examined 50 publications from each of the six journals Archives of Disease in Childhood (Arch Dis Child), Pediatric Research (Pediatr Res), Journal of Pediatrics (J Pediatr), The Lancet (Lancet), The BMJ (BMJ), and The New England Journal of Medicine (NEJM) published after June 1, 1997. (NEJM). From this, Albert (2016) developed the 'standard journal structure,' which he refers to as 2/7/7/6, which implies that the introduction should take up two paragraphs, the methods should take up seven, the findings should take up seven, and the commentary should take up six. Albert (1996) was a little more lenient, recommending a 2-3/4-6/4-6/5-8 arrangement. Other proposals based on Albert's empirical study, such as 2/5/5/4 for shorter pieces, have also been offered. Araujo (2014) looked at original papers from Arquivos Brasileiros de Cardiologia in January 2012 and 2013, and the first two issues of the Journal of the American College of Cardiology from the same years suggested 3/6-9/4-9/10 paragraphs. If then use the DerSimonian and Laird (DerSimonian & Laird, 1986) technique to meta-analyze the data supplied by Albert (2016) per section, the proposed structure is 3/8/7/7 (Table 1), resulting in a total of 25 paragraphs. However, Albert's work is over 20 years old and has not been updated.

Our article's initial goal is to present current findings on the structure of original papers published in major medical journals. Since the availability of reporting guidelines, such as the Schulz et al.,(2010) or Bossuyt et al., (2015) statement, has changed the way studies are reported over the last 20 years, we hypothesize that more recent journal articles will have an increasing number of paragraphs over time, particularly in the methods section. To this goal, we examine data from original publications published in the journals BMJ, Lancet, The Journal of the American Medical Association (JAMA), NEJM, and PLOS Medicine in 2005, 2010, and 2015. (PLOS). Because there are no page restrictions, we predict PLOS, an electronic-only publication, to include more paragraphs than print journals BMJ, JAMA, Lancet, and NEJM. Our second goal is to broaden the statistics to include additional structural characteristics such as the number of tables, figures, and references and the availability of extra content. We anticipate that newly published papers will feature other supplemental content.

Materials and methods

The number of paragraphs per section in original publications published in the English language medical journals BMJ, JAMA, Lancet, NEJM, and PLOS between 2005 and 2015. Researchers chose 30 papers randomly from each journal and year of publication, totalling 450 original research publications. They looked for the IMRAD structure, tallied the number of tables, figures, and references, and saw any supplemental material for each piece.

For continuous outcomes, means and SD were computed for each year and journal, and absolute and relative frequencies for the availability of extra content were reported. Random effect (RE) meta-analyses were performed using the DerSimonian and Laird (1986) technique, accounting for variation in the number of paragraphs between overtime and journals. These

analyses were carried out for each journal over time, each year, and for all years and journals combined. The standard errors and pooled RE estimates were determined. For the number of paragraphs and the % of journals providing supplemental material, linear mixed models with the journal as RE and year as a fixed effect (FE) were used to evaluate the influence of time. Estimates of the impact and their 95% confidence intervals were provided. An accurate mixed model with year as RE and journal as FE was used to examine the hypothesis that PLOS contains more paragraphs than print journals.

Multiple testing was not adjusted, and the significance threshold for all analyses was set to 0.05. R 3.6.2 was used to conduct all statistical analyses. As an add-on, data and analytic code written in Markdown are accessible.

Results

The researchers looked at 450 original research publications from five medical journals. In all of them, the IMRAD framework was employed. Over time, all journals have witnessed a rise in the number of paragraphs in each section and overall. JAMA and NEJM, on the other hand, have seen a tiny increase. BMJ and Lancet had a noticeable rise in the techniques section, with an average of 5 and 4 more paragraphs, respectively. PLOS also saw an increase in the number of paragraphs, with one, five, three, and two extra paragraphs in the introduction, methods, results, and discussion for 2015 and 2005, respectively.

While the overall number of figures and tables stayed relatively steady, the number of references climbed between 2005 and 2010 and 2015, with BMJ and PLOS having the most significant increases in connections.

The amount of original articles containing extra content has increased the greatest. In 2005, no JAMA publication included supplemental information; by 2015, 97 % (29 of 30) of JAMA articles featured supplementary content.

The overview (0.03 pars. per year, 95 percent CI: 0.00–0.06) and methods sections (0.29 pars. per year, 95 percent CI: 0.19–0.39) had the most minor and most significant increases in the number of paragraphs each year, respectively. In 2005, the typical article structure was 3/9/9/8, but by 2015, it had risen to 3/11/10/8. Within ten years, techniques had expanded by two paragraphs on average, while outcomes had increased by one section. This equates to a one-paragraph rise every two years (1.08 pars. every two years; 95 percent confidence interval: 0.70–1.45). The pooled standard article format is 3/10/9/8, taking all years and journals into account.

Discussion

The IMRAD framework was used in all the five main medical publications studied in all original papers. Between 2005 and 2015, the overall number of paragraphs climbed by one every two years, with the techniques section seeing the most growth. While the NEJM had the most paragraphs in the Albert analysis from 1996 (Albert, 1996), it had the fewest paragraphs on average across the years. PLOS contained the most sections, as predicted, and papers in PLOS were around 3.5 paragraphs longer than those in print journals. In addition, PLOS has the most figures and references per article. In comparison to 2005, all five of the journals studied now offer extra information.

The EQUATOR network—which stands for Enhancing the Quality and Transparency of Health Research—provides reporting requirements and checklists for a wide range of research and study kinds. Several publications demand that writers follow these standards to the letter.

The initial edition of the CONSORT statement (Consolidated Standard of Reporting Trials) was issued in 1996 and had 21 items.

Albert proposed using 2/7/7/6 paragraphs for the introduction/methods/results/discussion sections based on his empirical work over 25 years ago. Other writers (Albert, 1996; Araujo, 2014) were more lenient in their suggestions. We update the standard article format, and we now propose 3/10/9/8 paragraphs for each of the four main divisions. Thus, the total number of sections is around 30. As a result, a suggestion for the number of words each paragraph is made. JAMA, for example, has a 3000-word restriction. With 30 sections, the primary content has an average of 100 words for each paragraph. Because cardiology publications have a word count restriction of around 4000 words after references, Araujo (2014) proposed 130 words for them.

Albert (2016) and Araujo (2014) both made suggestions for the subjects of the various paragraphs. Their ideas for themes, however, are vastly different. According to Albert, the first paragraph of the discussion should quickly review the essential findings. However, Soares de Araujo believes that the research problem should be stated again in the first paragraph of the discussion. The repeating of the research problem, in our opinion, is unnecessary, and we agree with Albert's basic notion of article writing. He came up with 6 ideas for the discussion section: 1) critical results summary, 2) study shortcomings, 3) study strengths, 4) how it fits in the literature, 5) implications for future research and 6) implications for policy/treatment.

Our research did not look at the influence of article structure on citation frequency. Several studies (Falagas et al., 2013; Habibzadeh & Yadollahie, 2010; Hudson, 2014; Jacques & Sebire, 2010; Letchford et al., 2015; Paiva et al., 2012; Schreuder & Oosterveld, 2008) found a link between the length of the title and the frequency of citations.

Shorter titles receive more citations; thus, the lesson here is to keep the title brief. The relationship between the linguistic difficulty of the title, abstract, and main text and citation frequency has also been investigated (Fox & Burns, 2015; Goodman, 2000; Jacques & Sebire, 2010; Lu et al., 2019; Whissell, 1999). While there was no correlation between citation frequency and main text-linguistic complexity (Lu et al., 2019), top-ranked journals utilize simple language in their titles and abstracts (Whissell, 1999).

On the other hand, scientific writings are notoriously difficult to read (Albert, 2004; Hall, 2006). For scientific publications, for example, the Gunning fog index (Dubay, 2004), which considers sentence length and word complexity, is around 17 (Roberts et al., 1994; Weeks & Wallace, 2002; Yeung et al., 2018). Even though text complexity is lowered following peer review, texts remain much more complicated than everyday newspaper stories with a fog index of 12 (Grover et al., 2014). On the other hand, insurance plans are far more complex, with a fog index of around 20 (Albert, 2004).

The authors also examined the relationship between citation frequency, page length, reference count, and author recognition (Didegah & Thelwall, 2013; Hwang et al., 2019). More extensive articles and those with a higher number of references are likely to be review articles. With these ideas in mind, it's reasonable to assume that papers with more pages and references will have more citations.

A reviewer of our paper emphasized the necessity of looking at the impact of the media, which has been paying more attention to journal papers in recent years, as well as changes in journal methods and author instructions. Future studies should look at these issues.

Conclusion

To maximize their work's chances of publishing, authors should not simply adopt the traditional IMRAD format, and they should also consider the basic layout of their intended publication. Authors should employ 3, 10, 9 and 8 paragraphs for the introduction, methods, results, and discussion sections, respectively, if a journal-specific format is not available. Supplementary material has become the norm, and it should be utilized whenever necessary. Because online publications do not have page constraints, authors should be aware that print journals may have a different format than online-only journals. Finally, and most critically, the target journal's instructions to writers must be strictly followed.

References

- Albert, T. (1996). Publish and prosper. *BMJ*, 313(7070), S2-7070. <https://www.bmj.com/content/313/7070/S2-7070.pdf+html>
- Albert, T. (2004). Why are medical journals so badly written? *Medical Education*, 38(1), 6–8. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2923.2004.01733.x>
- Albert, T. (2016). *Winning the Publications Game: The smart way to write your paper and get it published*. CRC Press. https://books.google.com/books?hl=en&lr=&id=z8f1CwAAQBAJ&oi=fnd&pg=PP1&dq=Winning+the+Publications+Game:+The+Smart+Way+to+Write+Your+Paper+and+Get+It+Published.+4th+ed.+Boca+Raton:+CRC+Press%3B+2016.+140p+p.&ots=eHUyUbT_NA&sig=uNTIPFoEe5NrpDbP3AO-cCRBy68
- Albert, T. (2018). *Winning the Publications Game*. CRC Press. <https://doi.org/10.1201/9781315384153>
- Araujo, C. G. S. de. (2014). Detailing the writing of scientific manuscripts: 25-30 paragraphs. *Arquivos Brasileiros de Cardiologia*, 102. <https://www.scielo.br/j/abc/a/wZPrXSbYfp66JN9BSvq4tHR/?lang=en>
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C. W., Kressel, H. Y., Rifai, N., Golub, R. M., Altman, D. G., Hooft, L., Korevaar, D. A., Cohen, J. F., & STARD Group. (2015). STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Clinical Chemistry*, 61(12), 1446–1452. <https://doi.org/10.1373/clinchem.2015.246280>
- Brain, L. (1965). Structure of the scientific paper. *British Medical Journal*, 2(5466), 868–869. <https://doi.org/10.1136/bmj.2.5466.868>
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177–188. [https://doi.org/10.1016/0197-2456\(86\)90046-2](https://doi.org/10.1016/0197-2456(86)90046-2)

- Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, 7(4), 861–873. <https://doi.org/10.1016/j.joi.2013.08.006>
- Dubay, W. (2004). The Principles of Readability. *CA*, 92627949, 631–3309. https://www.researchgate.net/publication/228965813_The_Principles_of_Readability
- Falagas, M. E., Zarkali, A., Karageorgopoulos, D. E., Bardakas, V., & Mavros, M. N. (2013). The impact of article length on the number of future citations: a bibliometric analysis of general medicine journals. *PloS One*, 8(2), e49476. <https://doi.org/10.1371/journal.pone.0049476>
- Fathi M Sherif. (2013). *Getting your paper published: what do editors like and not like?* https://www.researchgate.net/post/Getting_your_paper_published_what_do_editors_like_and_not_like
- Fox, C. W., & Burns, C. S. (2015). The relationship between manuscript title structure and success: editorial decisions and citation performance for an ecological journal. *Ecology and Evolution*, 5(10), 1970–1980. <https://doi.org/10.1002/ece3.1480>
- Froese, F. J., & Bader, K. (2019). Surviving the desk-review. In *Asian Business & Management* (Vol. 18, Issue 1, pp. 1–5). Springer. <https://link.springer.com/article/10.1057/s41291-019-00060-8>
- Goodman, N. W. (2000). Survey of active verbs in the titles of clinical trial reports. *BMJ (Clinical Research Ed.)*, 320(7239), 914–915. <https://doi.org/10.1136/bmj.320.7239.914>
- Grover, V., Raman, R., & Stubblefield, A. (2014). What Affects Citation Counts in MIS Research Articles? An Empirical Investigation. *Communications of the Association for Information Systems*, 34. <https://doi.org/10.17705/1CAIS.03474>

- Habibzadeh, F., & Yadollahie, M. (2010). Are shorter article titles more attractive for citations? Cross-sectional study of 22 scientific journals. *Croatian Medical Journal*, 51(2), 165–170. <https://doi.org/10.3325/cmj.2010.51.165>
- Hall, J. C. (2006). The readability of original articles in surgical journals. *ANZ Journal of Surgery*, 76(1–2), 68–70. <https://doi.org/10.1111/j.1445-2197.2006.03651.x>
- Hill, A. B. (1965). Report of Editors' Conference. *British Medical Journal*, 2(5466), 870–872. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1846363>
- Hochberg, M., & Hochberg, M. E. (2019). *An Editor's Guide to Writing and Publishing Science*. Oxford University Press, USA. <https://books.google.com/books?hl=en&lr=&id=LCeeDwAAQBAJ&oi=fnd&pg=PP1&dq=1.+Hochberg+M.+An+Editor%27s+Guide+to+Writing+and+Publishing+Science.+Oxford:+Oxford+University+Press%3B+2019.&ots=hewpjSVVj9&sig=u0hqJQtRh-ZqUR0sInHFVWpuHf0>
- Hudson, J. (2014). An analysis of the titles of papers submitted to the UK REF in 2014: authors, disciplines, and stylistic details. *Scientometrics*, 109(2), 871–889. <https://doi.org/10.1007/s11192-016-2081-4>
- Hwang, A., Arbaugh, J. B., Bento, R. F., Asarta, C. J., & Fornaciari, C. J. (2019). What causes a Business and Management Education article to be cited Article author or journal. *The International Journal of Management Education*, 17(1), 139–150. <https://www.sciencedirect.com/science/article/pii/S147281171830051X>
- Jacques, T. S., & Sebire, N. J. (2010). The impact of article titles on citation hits: an analysis of general and specialist medical journals. *JRSM Short Reports*, 1(1), 2. <https://doi.org/10.1258/shorts.2009.100020>
- Korner, A. M. (2008). *Guide to publishing a scientific paper*. Routledge. <https://www.taylorfrancis.com/books/mono/10.4324/9780203938751/guide-publishing->

scientific-paper-ann-körner

- Letchford, A., Moat, H. S., & Preis, T. (2015). The advantage of short paper titles. *Royal Society Open Science*, 2(8), 150266. <https://doi.org/10.1098/rsos.150266>
- Lu, C., Bu, Y., Dong, X., Wang, J., Ding, Y., Larivire, V., Sugimoto, C. R., Paul, L., & Zhang, C. (2019). Analyzing linguistic complexity and scientific impact. *Journal of Informetrics*, 13(3), 817–829. <https://www.sciencedirect.com/science/article/pii/S175115771830227X>
- Paiva, C. E., Lima, J. P. da S. N., & Paiva, B. S. R. (2012). Articles with short titles describing the results are cited more often. *Clinics*, 67, 509–513. <https://www.scielo.br/j/clin/a/9v54CT4wzZ3VLQM4xrTGnGF/abstract/?format=html&stop=next&lang=en>
- Roberts, J. C., Fletcher, R. H., & Fletcher, S. W. (1994). Effects of peer review and editing on the readability of articles published in *Annals of Internal Medicine*. *JAMA*, 272(2), 119–121. <http://www.ncbi.nlm.nih.gov/pubmed/8015120>
- Schreuder, M. F., & Oosterveld, M. J. S. (2008). Who ever said size doesn't matter? The association between journal title length and impact factor. *NDT Plus*, 1(2), 126–127. <https://doi.org/10.1093/ndtplus/sfm047>
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials. *PLoS Medicine*, 7(3), e1000251. <https://doi.org/10.1371/journal.pmed.1000251>
- Sollaci, L. B., & Pereira, M. G. (2004). The introduction, methods, results, and discussion (IMRAD) structure: a fifty-year survey. *Journal of the Medical Library Association*, 92(3), 364. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC442179/>
- Thrower, P. (2013). Eight reasons I accepted your article. *Elsevier*, 1–3. <https://www.researchgate.net/file.PostFileLoader.html?id=55045237d11b8b482f8b463b&as>

setKey=AS%3A273733564993546%401442274575591

Weeks, W. B., & Wallace, A. E. (2002). Readability of British and American medical prose at the start of the 21st century. *BMJ (Clinical Research Ed.)*, 325(7378), 1451–1452. <https://doi.org/10.1136/bmj.325.7378.1451>

Whissell, C. (1999). Linguistic complexity of abstracts and titles in highly cited journals. *Perceptual and Motor Skills*, 88(1), 76–86. <https://journals.sagepub.com/doi/abs/10.2466/pms.1999.88.1.76>

Yeung, A. W. K., Goto, T. K., & Leung, W. K. (2018). Readability of the 100 Most-Cited Neuroimaging Papers Assessed by Common Readability Formulae. *Frontiers in Human Neuroscience*, 12, 308. <https://doi.org/10.3389/fnhum.2018.00308>