EasyChair Preprint

№ 7108

Assessment of Quality and Readability in the Observational COVID Studies from ELSEVIER Database: Global Literature on Coronavirus Disease 2020 -2021

Dairamise Rodas, Otylia Marion Nowicka,
Mohammed Alhamadani and Aakash Maroti

November 28, 2021

Assessment of Quality and readability in the observational COVID studies from ELSEVIER database: Global literature on coronavirus disease 2020 -2021

Dairamise B Rodas Martínez
Otylia Marion Nowicka
Mohammed ALHamadani
Aakash Maroti

Master Program: CBEM and MIIS
daramiseberlet.rodas01@estudiant.upf.edu
otyliamarion.nowicka01@estudiant.upf.edu
mohammed.alhamadani01@estudiant.upf.edu
aakash.maroti01@estudiant.upf.edu

Abstract. COVID19 has had a huge impact on worldwide health from its beginning, the new disease comes with new researches and literature, so it is highly important to know their quality and readability.

The panoramic view of a new disease is given at first by the observational type studies, that's why the aim of this study is to determine the percentage of accomplishment of each quality aspect, and determine the text readability from a well-known health dataset like ELSEVIER. We measured the quality of the ELSEVIER COVID19 observational articles by scoring its different characteristics (clarity of aims, relevance, sources, originality). Also, to measure their readability: the Flesch Kincaid reading ease tool was used.

The results showed that the measured ELSEVIER COVID 19 observational articles have a 4.5 quality score points, achieving a 90% of research quality in total. The mean Readability Score was 34.31, indicating a hard legibility, that is expected for scientific articles.

Keywords: Quality Score, Readability, Evaluation Scientific papers, Observational COVID19 studies.

1 Introduction

Quality in research papers

A research paper's quality is determined by the research project it reports on. There is, however, a lot that authors may do to improve the clarity and utility of their papers. The guidelines for authors in journals frequently focus on the structure, style, and length of publications, but they don't always emphasize the importance of properly explaining the science work and ethics, so this review reminds researchers that transparency is important. The research question should be stated clearly and explain where it came from and why it is essential. The research methods must be fully described and, when applicable, in accordance with evidence-based information and reports such as the CONSORT statement for randomized controlled trials. If the study was a trial, the publication should explain where, when, and how the trial was

registered, as well as the registration number. Finally, any potential conflicts of interest must be disclosed. [1]

The aim of the research is to Assess the Quality and readability of the observational COVID studies from the ELSEVIER database. Elsevier is a Dutch-based publishing house that specializes in scientific, technical, and medical publications. It is a member of the RELX Group, which was previously known as Reed Elsevier until 2015. Journals like The Lancet and Cell, the ScienceDirect collection of e-journals, the Trends and Current Opinion series of journals, the online citation database Scopus, the SciVal tool for measuring research performance, the ClinicalKey search engine for clinicians, and the ClinicalPath evidence-based cancer care service are among the company's products. Digital solutions for data administration, training, research analytics, and assessment are among Elsevier's products and services. Every year, Elsevier publishes about 500,000 papers in 2,500 publications. It has about 17 million documents and 40,000 e-books in its archives. More than 1 billion downloads are made each year.[2]

The quality is measured by the discerning, which is a tool or instrument designed to assist users of consumer health information in evaluating the quality of written information about treatment options. And using the Flesch Reading Ease to measure the readability that assigns a score to a text ranging from 1 to 100, with 100 being the most readable. A score of 70 to 80 corresponds to a school grade level of 8. This implies that literature should be relatively simple to read for the ordinary adult. Usually, the scientific papers have a score lower than 40 due to their technical and complexity. The test works by taking into account sentence and word counts. The mathematical formula underlying it looks like this: 20.835 - 1.015 (total words/ total sentences) - 84.6 (total syllables/ total words). It is convenient to use online tools to measure this legibility score.[3]
There exist several types of investigation evaluations systems, one of them is the DISCERN test, which is a probed health articles assessment instrument, consists of 15 questions plus an overall quality rating.[4]. The present study uses it like a template, formulating the quality questions to represent a separate quality criterion – an important feature or standard that is an vital part of good quality information on treatment choices. And these questions represent the characteristic that the research will be evaluated by them:

The clearance of the aim: A high-quality publication will have well-defined aims. An overview of a publication should begin with a description of what it is about, what it covers, and whom it is intended for. Clear goals at the start of a publication are important because they show what aspects of the problem and its treatment will be covered, and they help you decide whether the publication will provide the information you need. It's especially important to understand what's not provided, as you may require extra information before making a treatment decision.

Relevance: A high-quality publication will focus on the needs of its readers. It's important that the information you receive about a possible treatment or possibilities is tailored to your lifestyle and circumstances. The publication should not make

unrealistic recommendations or contain assumptions or language that you find insulting or improper.

The reality in sources: Possible treatments information should be accurate and based on the best scientific evidence available. Because this would necessitate verification against additional sources, DISCERN cannot be used to determine whether the information is factual or based on sound evidence. A high-quality publication, on the other hand, will make it clear where the evidence for treatment options came from. Details about the sources of evidence are important because they allow you to double-check the information or decide whether you need to look for more.

Originality: A publication should be truthful and educational. It should not sway your opinion by 'pushing' specific treatment options or employing shock tactics.' A high-quality publication will include a reference to 'grey' areas where the most effective treatment is unknown. This ambiguity could be due to the following factors:
1- There is no evidence about successful possible treatments;
2- The evidence that does exist is contradictory.
3- It's unclear who is most likely to benefit or be harmed as a result of the therapy option.[4]

2 Research methodology

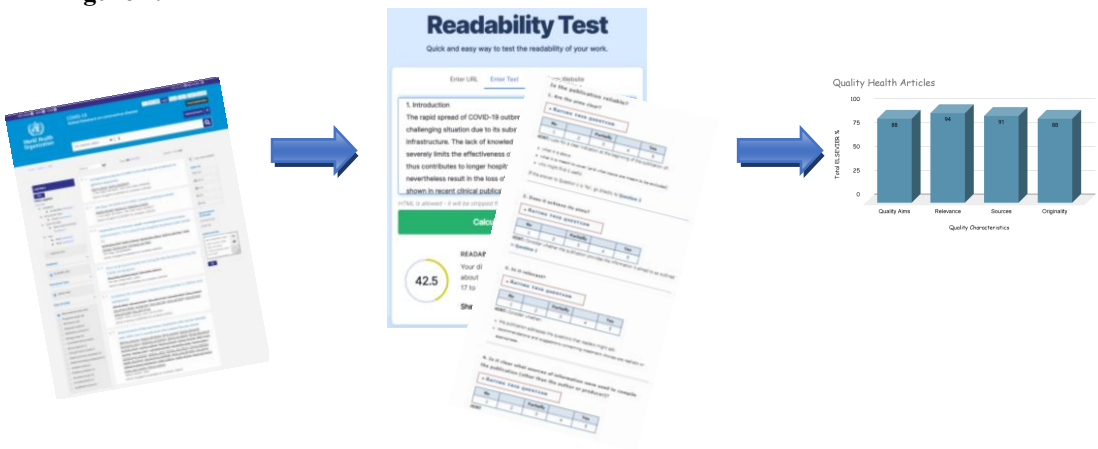
In this observational descriptive study. Our sample n=40 was taken from the World Health Organization research engine for Global literature coronavirus disease (with the filters: dataset: ELSEVIER, document type: article, type of study: observational, year: 2020 and 2021. [5] (See the workflow in figure 2.1) The whole ELSEVIER's observational articles found were N= 104. Over the sample was applied a questionnaire (Table 2.1) that we developed from diverse quality health research assessment instruments [4, 6, 7] fig 2.1. Also, in each article was measured a Flesch Kincaid reading ease, an English text readability score (Table 2.2) [8]. With the obtained data was calculated the percentage of each quality characteristic and the legibility of the papers respectively.

We consider that the chosen methodology is the most suitable because the purposes of the study are merely descriptive, like indicating our research question and aims. [9, 10] Another methodology wouldn't suit our research aims, for instance: Not a theoretical inductive/ deductive study because we are not making a state of art article, we are quantitatively measuring some qualitative aspects, but not searching for a theory, model or new method. A historical method is not the best option due to the limitation on time, we are not making projections to the future. Either an experimental method because we are not intervening in any variable to see what would be the result. [6]

The strengths of transversal descriptive research are: don't need large amounts of money, are relatively quick to realize, it is reproducible, serves like a good characterization tool and the first step to develop further analytic studies, useful in the planification and sanitary administration, it serves like an overview of new problems.

The limitations of the chosen methodology are: Do not permit incidence calculation, do not estimate the risk of a determinative factor, not useful to infrequent diseases, and do not inform about real association between variables. [11]

Figure 2.1



Recollecting Data **Calculating Scores** **Obtaining results**

Table 2.1 Questionnaire: measurement of quality in health articles.

Code	Quality question	Measured characteristic
Q1	Are the aims clear?	Quality of Research Aims
Q2	Does it achieve its aims?	
Q3	Is it relevant?	Relevance
Q4	Is it clear what sources of information were used to compile the publication (other than the author or producer)?	Quality in sources
Q5	Is it clear when the information used or reported in the publication was produced?	
Q6	Is it balanced and unbiased?	Originality
Q7	Does it provide details of additional sources of support and information?	
Q8	Does it refer to areas of uncertainty?	

See appendix 1 to know the hints that were looked for in order to respond to the questions.

Table 2.2 Flesch Kincaid reading ease

Score	School level (US)	Notes
100–90	5th grade	Very easy to read. Easily understood by an average 11-year-old student.
90–80	6th grade	Easy to read. Conversational English for consumers.
80–70	7th grade	Fairly easy to read.
70–60	8th & 9th grade	Plain English. Easily understood by 13- to 15-year-old students.
60–50	10th to 12th grade	Fairly difficult to read.
50–30	College	Difficult to read.
30–10	College graduate	Very difficult to read. Best understood by university graduates.

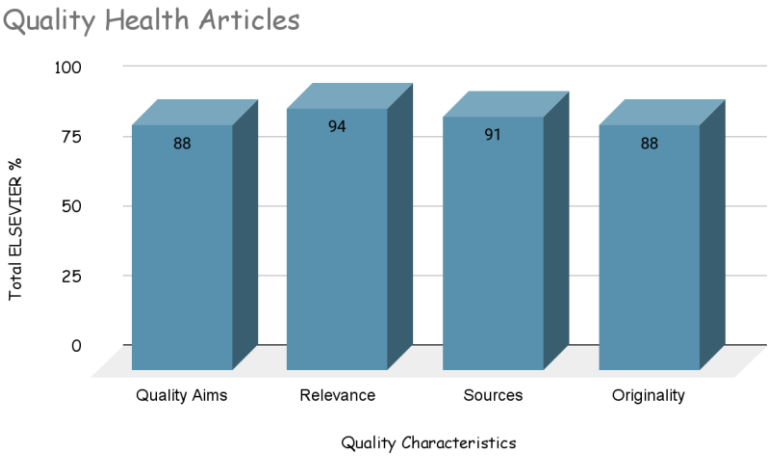
3 Results

Table 3.1 Quality Health Papers

Question	Main Score	Total Score	Quality Characteristics	Total Score
Q1	4.5	90.77%	Clarity Aims	88.97%
Q2	4.3	87.18%		
Q3	4.7	94.59%	Relevance	94.59%
Q4	4.7	93.85%	Quality in Sources	91.92%
Q5	4.5	90.00%		
Q6	4.7	93.33%		
Q7	4.1	82.05%		
Q8	4.4	87.69%	Originality	88.84%

All the main scores measured with the questionnaire were higher than 4, achieving 88.97% in clarity aims, 94.59% in relevance, 91.92% in quality in sources and 88.84% in self peer review.

Figure 3.1 The total ELSEVIER in (%) vs the quality characteristics



All the main scores measured with the questionnaire were higher than 4, achieving 88.97% in clarity aims, 94.59% in relevance, 91.92% in quality in sources and 88.84% in self peer review.

Table 3.2 Readability Score of ELSEVIER health COVID-19 articles

Readability SCORE	Total COVID Articles
< 40	27
> 40	13

The majority of the articles that were analyzed had low readability scores. This means that these articles were written with complex vocabulary and sentence structure.

4 Conclusions

Suitable assessment of the quality in the COVID19 observational research [12] is essential to the well interpretation of primary research and to conduct reviews and meta analysis.

We found that the ELSEVIER dataset accomplished a high qualification: 4.5 or a 90%, for the developed health papers assessment instrument. Since there is no publicly available information or statistics on comparison of scientific publishing companies, this result provides a useful insight.

As it was expected that the readability Flesch Kincaid reading ease score was low: 34.31 In this study for ELSEVIER Covid19 observational articles. And the decreasing readability of scientific articles is already a well documented issue [13]. The usual range for scientific papers is found to be between 0 - 50 according to the study.[14] Which is legible to the scientific community but not so much for the non-science lectors.

We recommend that the media, channels and social media's science disseminators that transform scientific information into a more understandable and entertaining way, should take more relevance and a respectable position in the scientific community, because in these times, when fake data is spread with so much facility, the truth should have a friendly way to reach the non-scientific population, in order to avoid misinformation and the severe consequences that it has in public health.

We encourage constant evaluations of the used datasets in order to have a better comprehension of what the researchers should see when they are making investigations.[15] It also reveals certain societal biases that might not be expected, such as, COVID-19 medical papers have fewer women first authors than expected. [16]

References

1. Groves, T. (2010). What makes a high-quality clinical research paper? In Oral Diseases (Vol. 16, Issue 4, pp. 313–315). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1601-0825.2010.01663.x
2. About ELsevier. About Elsevier. (n.d.). Retrieved November 18, 2021, from https://www.elsevier.com/about.
3. Readable. 2021. Flesch Reading Ease and the Flesch Kincaid Grade Level – Readable. [online] Available at: <https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level/> [Accessed 19 November 2021].
4. Bc. Charnock, D. (1998). The DISCERN handbook. Quality criteria for consumer health information on treatment choices. Radcliffe: University of Oxford and The British Library
5. WHO COVID19 Global literature in coronavirus disease, https://search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/?lang=en

6. Charnock, D. (1998). The DISCERN handbook. Quality criteria for consumer health information on treatment choices. Radcliffe: University of Oxford and The British Library.
7. Kumar, R., (2011) Research Methodology. A step-by-step guide for beginners (3rd Edition), Sage Publications.
8. Fx readability tool, https://www.webfx.com/tools/read-able/
9. Goddard, W., Melville, S.: Research Methodology: An introduction. Juta y Co. Ltd, Lansdowne (2001) https://books.google.com/books?id=bJQJpsU2a10C&lpg=PA1&ots=XuqbO9zP9n&dq=research%20methodology&lr&hl=es&pg=PA8#v=onepage&q=research%20methodology&f=false
10. Portell, M., Anguera, M. T., Chacón-Moscoso, S., & Sanduvete-Chaves, S. (2015). Guidelines for reporting evaluations based on observational methodology. Psicothema, 27(3), 283–289. https://doi.org/10.7334/psicothema2014.276
11. Salinero, J. G. (n.d.). Estudios descriptivos. http://webpersonal.uma.es/de/jmpaez/websci/BLOQUEIII/DocbIII/Estudios%20descriptivos.pdf
12. Simon Sanderson, Iain D Tatt, Julian PT Higgins, Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography, International Journal of Epidemiology, Volume 36, Issue 3, June 2007, Pages 666–676, https://doi.org/10.1093/ije/dym018
13. Research: The readability of scientific texts is decreasing. eLife 2017;6:e27725. time.https://elifesciences.org/articles/27725
14. Readable. 2021. Flesch Reading Ease and the Flesch Kincaid Grade Level – [online] Available at: <https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level/> [Accessed 19 November 2021]
15. Ioannidis, J. P. A. (2018). Meta-research: Why research on research matters. PLoS Biology, 16(3). https://doi.org/10.1371/journal.pbio.2005468
16. Meta-Research: COVID-19 medical papers have fewer women first authors than expected. https://elifesciences.org/articles/58807 eLife 2020;9:e5880714. Ioannidis, J. P. A. (2018). Meta-research: Why research on research matters. PLoS Biology, 16(3). https://doi.org/10.1371/journal.pbio.2005468

Appendix 1

1. Are the aims clear?
2. Does it achieve its aims?
3. Is it relevant? (the publication addresses the questions that readers might ask. Recommendations and suggestons are realistic or apropiated we can find it in the discussion and abstract)

4. Is it clear what sources of information were used to compile the publication (other than the author or producer)? Check whether the main claims are accompanied by a reference to the sources used as evidence Look for a means of checking the sources used such as a bibliography/reference list or the addresses of the experts or organisations quoted, or external links to the online sources.
5. Is it clear when the information used or reported in the publication was produced? dates of the main sources of information used to compile the publication date of any revisions of the publication date of publication
6. Is it balanced and unbiased? (Look for a clear indication of whether the publication is written from a personal or objective point of view. Be wary if isn't presented in a sensational, emotive or alarmist way)
7. Does it provide details of additional sources of support and information? Look for suggestions for further reading or for details of other organizations providing advice and information about,
8. Does it refer to areas of uncertainty? Look for discussion of the gaps in knowledge or differences in expert opinion