

Wikipedia as a health information resource in various contexts

Denise A. Smith

McMaster University, Health Sciences Library

Western University, Faculty of Information and Media Studies

What we already know

English language medical content on Wikipedia has had **more page views** than any other health information resource online

Two broad comprehensive reviews of Wikipedia in academic literature have already been done:

Okoli C., Mehdi M, Mesgari M., Nielson F.A., Lanamäki A. [Wikipedia in the eyes of its beholders: A systematic review of scholarly research on Wikipedia readers and readership.](#) J Assoc Inf Sci Technol. 2014 Dec;(12):2381.

Mesgari M, Okoli C, Mehdi M, Nielsen FÅ, Lanamäki A. ["The sum of all human knowledge": A systematic review of scholarly research on the content of Wikipedia.](#) J Assoc Inf Sci Technol. 2015 Feb;66(2):219–45.

When it comes to Wikipedia, we already know that it's the most frequently accessed online resource for health information: its English language medical content has received more unique pageviews than any other health information resource online (3). We also know that people turn to the internet to look for health information. This was established in a 2013 paper by James Heilman (Doc_James) and Andrew West. Efforts to *understand*, in depth, Wikipedia's role in the landscape of health information resources are limited. Two comprehensive literature reviews of the scholarly literature covering Wikipedia have been conducted: one that focuses on readers and readership and another that looks at Wikipedia's content, but the academic literature and its treatment of Wikipedia as a health information resource hadn't yet been reported.

What we don't know yet

How do we reconcile the fact that Wikipedia's medical content **is** used with other published evidence?

Is the medical content good?

Is the medical content used in different contexts? What are they?

But how do we reconcile what we know to be true, that Wikipedia's medical content is being used, with what there is no consensus about: whether it's good enough, how it's best used, in what contexts it's being viewed. When Wikipedians advocate to health professionals and experts to contribute to Wikipedia's medical content, the only leverage we really have that's backed up with evidence is that the medical content is being looked at. But we haven't really used the breadth of the literature on Wikipedia in a health context to paint a clear picture of all the various contexts in which Wikipedia is employed as a health information source.

So that's what I wanted to do.

Research Questions

RQ1: What publication trends exist, if any, exist among health-related articles about Wikipedia

RQ2: What does the existing body of literature say about Wikipedia as a health information resource?

So I decided to do a comprehensive literature review of all the academic literature that had been published about Wikipedia within any health context.

Two research questions:

1. *What publication trends exist, if any, among health-related articles about Wikipedia?*

I was specifically interested in learning whether these academic articles would be published using an open access model and whether any trends emerged in the number of relevant publications over the years

2. *What does the existing body of literature say about Wikipedia as a health information resource?*

What I learned...

Publication Trends:

1. Open access articles about Wikipedia in the health context proportionately greater than open access articles in academic literature at large
2. Overall upward trend in the publication of health literature about Wikipedia

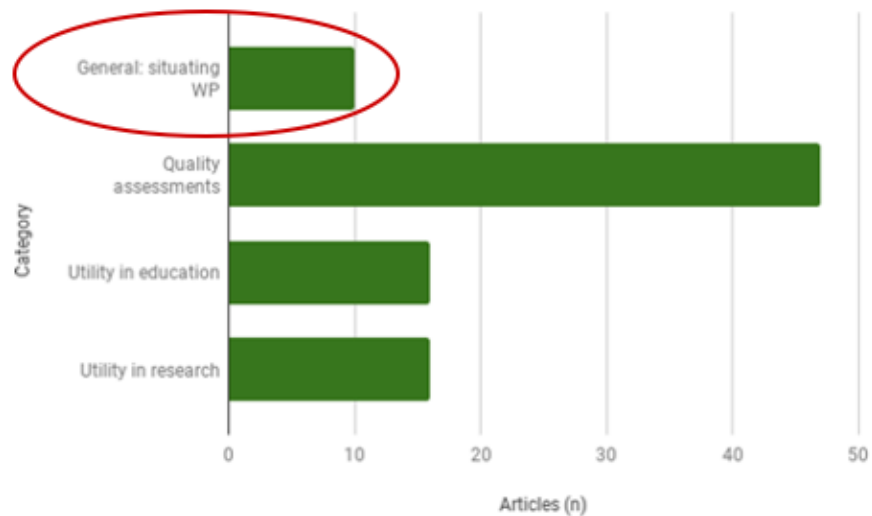
And I learned quite a lot.

With respect to publication trends I learned two things of interest:

1. A 2018 study estimated that approximately 28% of all literature is published using an open access publication model. Comparatively, of the 89 articles I included in my study 40% we published using the open access model. So, while it's still more common for articles about Wikipedia in a health context to be published in in subscription journals, the complement of OA literature within this review is actually significantly greater than the complement of OA literature in the larger body of all academic literature. This could be interpreted as an indicator that the authors of these studies, like Wikipedia, value open and broad dissemination of knowledge and this was encouraging to me.

2. I found an overall upward trend in the production of health literature about Wikipedia. The growth of publications in this field could be a promising indicator of the growth in the acceptance of Wikipedia within the academic health community.

Themes in the literature



The focus of my talk today is going to be primarily on the themes that emerged from the collected of 89 articles I managed to curate for this study. I might go into some detail, but the paper goes into extensive detail and even offers tables that outline which of the 89 articles fell into each theme that emerged. So if you're keen on learning the finer details I really encourage you to take a look at the paper in Plos One if you can.

Four major themes emerged as I engaged with the literature during this review:

1. Articles that focused on establishing Wikipedia as a source of health information.
2. Articles that assess the quality of Wikipedia's health and medical content

3. The utility of Wikipedia's health and medical content in education
4. The utility of Wikipedia's health and medical content for research

****CLICK FOR ANIMATION****

The articles that fall into the first them, situating Wikipedia as a health information source, are generally heavily cited articles that position Wikipedia as a health information resource by using its usage statistics, the superiority of Wikipedia's medical content as compared to non-specific general content in Wikipedia, the ubiquity of Wikipedia's medical content in the results of search engine queries, the sheer magnitude of the medical content in Wikipedia, and the vast numbers (+950,000) of references to high-quality publications in its pages, including strong representation from top-tier medical journals (NEJM, Lancet, Nature).

But by and large the literature on Wikipedia's health content was dominated by assessments of the quality of its medical pages. Articles of this nature accounted for more than half the articles reviewed in this study and within those assessments, even more contexts emerged.

Themes in the literature

Assessments of the quality of Wikipedia's medical content for:

- Consumers
- Students
- Professionals
- General assessments (no population or context explicitly identified)

QUALITY OF WIKIPEDIA'S MEDICAL CONTENT

Assessments of Wikipedia's quality could be categorized by the population of concern to the assessment: a general information consumer or patient, students, health professionals, and in some of the articles a specific user group wasn't clearly identified. In this study, I considered any article a quality assessment if it evaluated readability, accuracy, currency, comprehensiveness, completeness of coverage, or number or quality of references in one or a group of Wikipedia entries..

For consumers

So, with the general consumer or patient in mind, there was a general consensus that Wikipedia is a suitable resource for patients. Interestingly, a 2010 study found that

while Wikipedia wasn't necessarily considered the superior resource, it was still the most preferred source. This in itself is a goldmine of opportunity for future research. Unfortunately though, these assessments of suitability for patients are facing a degree of irrelevance because even the most recent published report is from 2013 and most of the evaluations focus on a specific health speciality or issue, so the findings can't be generalized.

For students

For students, assessments of the quality of Wikipedia's health content brought up two salient points: Wikipedia is used regularly by medical students (and we already knew this) and there is absolutely no consensus on its suitability for this population. Some of the articles conclude Wikipedia is suitable for medical students, others don't, and the findings are entirely contingent on the context of the evaluation, whether a comparator information source was used and if so, what Wikipedia was compared to. Some studies base their finds on comparing Wikipedia with other health information, like a textbook, which is not an appropriate comparator. With the textbook example, Wikipedia doesn't claim to be equivalent to a textbook for vascular surgery so it shouldn't really be compared to one.

For professionals

Again - another area where the quality of Wikipedia for a specific population, in this case health professionals, is brought into question. And again, there is no consensus among the studies. The same issues with comparisons to inappropriate information

sources also come up among these studies. Of course Wikipedia is not going to be as comprehensive as a tool for pharmaceutical medicines. We know that Wikipedia intentionally does not summarize information on dosage recommendations for example, out of concern for public safety. This would not be a concern for a proprietary database intended exclusively for pharmacists.

Quality of Wikipedia in all user contexts

The studies that had no specific audience in mind largely aligned with those that did. Some drew inappropriate comparisons, while others focus their comparisons of Wikipedia to other open, free tertiary resources or used a validated instrument, such as the DISCERN tool to determine the quality of Wikipedia's health content.

Themes in the literature

Assessments of the quality of Wikipedia's medical content for:

- Poor use of like-like comparison
- Readability of medical content remains a challenge
- Studies of quality on Wikipedia are futile, at best (unless they are used to make targeted, direct improvements on a specific article or collection of articles)

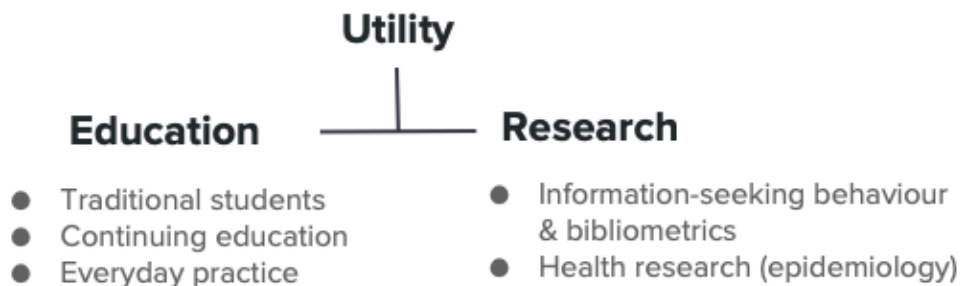
The biggest challenge for all of English Wikipedia's health content is that while it aims to be written in plain, accessible English, the readability of its medical content is often too low to accommodate the very people it stands to benefit most. The lowest reading level reported is around ninth grade and some studies find that Wikipedia's medical content requires at least a college reading level or is considered highly difficult. Other studies simply report Wikipedia's readability as higher than other health information sources like WebMD, MedlinePlus, or MayoClinic. While there's some evidence to show that this is getting better, there is still a lot of work to do to make Wikipedia's health information easy to read for all who access it.

Perhaps the most significant critique I have for these quality assessments is that that

wrongly attempt to generalize their findings to generate a holistic evaluation of all of Wikipedia's medical content. This is futile, in my opinion, and represents a broader misunderstanding of what Wikipedia is.

Wikipedia's articles are individual pieces of a larger whole, creating a mosaic of information where each piece contributes to the summarization of medical knowledge but where some pieces are more complete than others. Wikipedia is also fluid. While it is prudent to describe Wikipedia as a tertiary information resource, it is not equivalent to a published encyclopedia. Not only is it incomplete, it is also dynamic: evolving, and expanding. Any evaluation will remain a snapshot in time that may or may not be applicable by the time the study hits the press. Unless these studies are used to establish a plan to improve Wikipedia's content, it was a challenge for me to understand their value.

Themes in the literature



UTILITY IN EDUCATION

Wikipedia's utility in education is also considered within the context of different populations, ranging from its general usage in educational settings to its deliberate inclusion in course or program curricula, specifically as an educational tool for evidence-based practice or information literacy, or both. I used the term "education" to apply to students enrolled in a formal medical or health education program at the graduate or undergraduate level, professionals enrolled in a continuing education program, workshop, or course, and professionals who use Wikipedia to inform their everyday practice. Wikipedia's medical content is reported as an excellent learning tool when students edit Wikipedia for course credit. It's also reported as heavily used among students of health and medicine. As an educational tool, it's reported utility is

in knowledge sharing between professionals or for daily reference in everyday practice.

UTILITY IN RESEARCH

With respect to Wikipedia's utility in research, two contexts emerge and they both employ Wikipedia's usage data: information seeking behaviour and bibliometric research, and health research, namely, epidemiological studies. Bibliometric studies have been used to assess whether Wikipedia is cited in academic literature and its influence on the popularity of, or access to, high-impact journal articles.

Epidemiological studies have either used or explored the utility of Wikipedia's daily pageview statistics and raw usage data for estimating burden of disease, to predict possible disease outbreaks, or to monitor seasonal illnesses such as the flu.

Implications

- We now have **documentation** that there is a strong body of evidence for Wikipedia's prominence as a health information resource
- Provides foundation to support future research

Thank you!

Plos One paper: <https://doi.org/10.1371/journal.pone.0228786>

So what does this all mean? Well, it means that with the publication of this review, there is -at the very least- documentation that the literature provides strong evidence to position Wikipedia as a prominent health information resource for the public, patients, students, and practitioners who are using the internet for health information. We know that Wikipedia is gaining popularity as a tool for education in the classroom and that there is potential for the usage data of Wikipedia's health content to contribute to the study of disease outbreaks. I would to hope that my critiques of the disproportionate attention paid to the quality of Wikipedia's health content will encourage readers to stop obsessing over whether Wikipedia is good enough and to just try to make it better, hopefully with a focus on making it easier to read for the layperson. Rather than rate the quality of Wikipedia within a specific domain and

recommend its use or discontinuance of use as an information resource, these findings could instead be utilized to as evidence to support advocacy for expert participation in the production of Wikipedia's content. Engagement with Wikipedia at the professional *and* academic level are crucial in order to make it better. Now that I am aware of all the various contexts in how Wikipedia's medical content is used, I intend to conduct further research into the experiences of the layperson. I hope this research will provide more depth in our understanding of why Wikipedia is used and if it's impacting decision making of the layperson.