‘WP2Cochrane’, a tool linking Wikipedia to the Cochrane Library: Results of a bibliometric analysis evaluating article quality and importance

Arash Joorabchi
Cailbhe Doherty
University College Dublin, Dublin, Ireland

Jennifer Dawson
Cochrane, UK

Abstract
Medical information on English Wikipedia was accessed over 2 billion times in 2018. Our goal was to develop an automated system to assist Wikipedia volunteers to improve articles with high-quality sources from journals such as The Cochrane Library. We created an automated indexing system by linking available reviews from the Cochrane library with disease-related Wikipedia articles and evaluating the relationship between the quality and importance of these articles with the number of relevant and cited Cochrane reviews. We first conducted a bibliometric analysis, identifying disease-related Wikipedia articles and relevant/cited Cochrane reviews. Citations were thematically coded, and descriptive statistics were calculated. Finally, separate multinomial logistic regression analyses were conducted for article quality and importance. The indexing system identified 4381 disease-related Wikipedia articles, 1193 (27%) of which cited a Cochrane review. Higher quality Wikipedia articles were more likely to cite a Cochrane review (p = 0.002), while lower quality articles were less likely to cite a Cochrane review (p < 0.0005). A greater number of Cochrane reviews are available for more ‘important’ Wikipedia articles (p < 0.005), and these articles were more likely to cite a Cochrane review (p < 0.005). This approach to an indexing system can be leveraged by Wikipedia contributors and editors seeking to update disease-related Wikipedia articles with relevant Cochrane reviews (thus improving their quality), and online information seekers in need of additional information to supplement their Wikipedia search.

Corresponding author:
Cailbhe Doherty, Insight Centre for Data Analytics, O’Brien Centre for Science, University College Dublin, Belfield, Dublin 4, Ireland.
Email: Cailbhe.doherty@insight-centre.org

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Keywords
Bibliometrics [MeSH], Encyclopedias as Topic [MeSH], Internet, knowledge translation, open knowledge, peer review, Wikipedia

Introduction

An increasing proportion of healthcare professionals have spent their entire lives in an environment where smartphones, streaming and web-surfing have been ubiquitous. These ‘digital natives’ expect web-access to be universal and information to be perpetually accessible. The way information is consumed and implemented in the healthcare setting is changing rapidly as increasing numbers of individuals occupy and migrate to the digital space.

Healthcare professionals experience a theoretical unknown in 50 per cent of their clinical encounters, while patients can experience thousands of questions over the timespan of a disease. A typical avenue to fulfilling these ‘information needs’ is the Internet; 280 million health-related queries are made on search engines like Google every day.

Wikipedia is among the top search results on Google. Created in 2001, Wikipedia is a multilingual, web-based, free content encyclopaedia project supported by the Wikimedia Foundation and based on a model of openly editable content. Wikipedia is the fifth most popular site on the web with 2,125,792 visits per day and is the most popular online healthcare resource globally. Indeed, the use of Wikipedia among students, physicians, pharmacists and nurses to fulfil their information needs is widespread. Medical-related Wikipedia articles are accessed over 10 million times per day across all languages. This, despite several groups advocating caution when relying on Wikipedia as a primary information resource. These cautionary notes notwithstanding, the reputation of Wikipedia in academia continues to improve, as does the perception of its quality.

Medical-related article development on Wikipedia is supported by volunteers, many of whom are affiliated with the WikiProject Medicine Foundation. Efforts to improve the quality of health-related articles on Wikipedia are numerous and include initiatives to encourage medical professionals and students to contribute to Wikipedia, translation efforts, partnerships with health organisations and institutions, and developing offline content for parts of the world that do not have access to the Internet. Cochrane began working with WikiProject medicine in 2014 to improve the quality and content of the health evidence available on Wikipedia. This partnership is intended to support the inclusion of evidence produced and shared by Cochrane, where appropriate, and the continued update of Wikipedia articles when new reviews are published.

Cochrane Reviews are systematic reviews of primary research in human health care and health policy, and are internationally recognised as representing a high standard in evidence-informed health care. Wikipedia has a precedent of sharing information from Cochrane Reviews to improve the quality of its health-related articles. Currently, over 3000 Cochrane systematic reviews are cited in Wikipedia articles. One of the difficulties of maintaining congruence between medical-related Wikipedia articles and available Cochrane reviews is the continually evolving evidence landscape. English Wikipedia contains over 32,000 medical-related articles and there are more than 7000 active Cochrane reviews published in The Cochrane Library. In addition, approximately 30 new Cochrane reviews and 30 updated reviews are published each month. Ensuring that the evidence being shared is presented accurately and appropriately and maintaining congruence between these two data sets represents a significant burden for a global community of Wikipedians. The inevitable discord between available Cochrane reviews and their citation on Wikipedia represents a significant opportunity for computer-automation technologies to harmonise content from a
high-quality evidence resource (The Cochrane Library) with one of the most popular online knowledge bases (Wikipedia).

The primary aim of this investigation was to develop a free, automated indexing system that links reviews from the Cochrane Library with health-related Wikipedia articles that can be used by online information seekers with minimal computational burden. Using this system, our objectives were as follows:

1. To index health-related Wikipedia articles for which Cochrane reviews were available (‘relevant’ Cochrane reviews), to determine whether those reviews were cited and, to verify if the cited reviews are the latest version available or need updating.
2. To identify important add in health-related Wikipedia articles for which no Cochrane reviews were available.
3. To evaluate the relationship between Wikipedia article quality, importance and its citation of relevant Cochrane review articles.
4. This hypothesis-confirming research seeks to corroborate Wikipedia’s own guidelines for article quality, which specifies that higher quality articles must have inline citations from reliable sources.39

**Methods**

**Data collection**

Between January and April 2019, a system of search implementation to identify disease-related Wikipedia articles and Cochrane reviews that were relevant to those articles was developed. This tool, named WP2Cochrane, is now publicly available, and is presented as an interactive Jupyter Notebook (https://mybinder.org/v2/gh/ajoorabchi/WP2Cochrane/master?urlpath=lab/tree/index.ipynb) that serves as a tutorial for the process of linking disease-related Wikipedia articles with relevant Cochrane reviews. The Jupyter Notebook pipeline provides scripts that process the raw data into tabularised results that can enable others to replicate and validate our analyses. The linkage results in HTML and CSV format and the Python source code are available on the project’s GitHub repository (https://github.com/ajoorabchi/WP2Cochrane). For non-technical users, a browser extension was developed for Google Chrome containing the Cochrane2WP tool (https://chrome.google.com/webstore/detail/wikipedia-%20cochrane-libr/cehpfefpnicpmejgidpkgeenapnfcakm; Figure 1). The extension activates when the user is browsing Wikipedia. While browsing a disease-related Wikipedia article, users are provided with a list of related Cochrane reviews by clicking on the extension’s icon. The extension retrieves mapping data from the Cochrane2WP’s repository on GitHub and, therefore, its results are updated on a monthly basis to accommodate newly published Cochrane reviews and new disease articles published on Wikipedia. The Cochrane2WP script is currently hosted on a Google cloud server and set to run once a month. The result of each new run will be automatically uploaded to the tool’s GitHub repository, and the historical results will be kept in a persistent storage folder for future trend analyses (https://github.com/ajoorabchi/WP2Cochrane/tree/master/persistent_storage).

The process of Wikipedia to Cochrane library linkage commences by first retrieving the list of disease-related Wikipedia articles as indexed on Wikidata, using an SPARQL query (https://w.wiki/3kg). Wikidata is a machine-readable knowledge base and a central storage of structured data of Wikipedia.40 Using the Wikidata Query Service (https://query.wikidata.org/), all entities on Wikidata that have a valid corresponding Wikipedia article and are identified as an instance of a ‘disease’, were retrieved (e.g. dyslexia; https://www.wikidata.org/wiki/Q132971). As of 30 April
In the second stage of the process, a search on PubMed is automatically conducted to identify an entire library of published Cochrane reviews for each of the 4381 disease-related articles on Wikipedia. PubMed currently indexes all reviews published in The Cochrane library. By combining Boolean operators, we conducted a search for each disease-related article in the aggregated list, using the article title and its reroute terms. Due to the potential use of terminological derivations for the same medical concept on Wikipedia and Cochrane (e.g. ‘sprained ankle’ vs ‘ankle sprain’), a list of abbreviations was parsed from the Medical Search Headings (MeSH) list along with text simplification (lower case, no punctuation) to increase the reliability of the method.

**Data extraction and coding**

The quality and importance of all disease-related Wikipedia articles (as judged by the articles’ editors) were first extracted. Wikipedia article quality is measured on a grade scale with nine levels (in order of descending quality: featured article, A-class article, good article, B-class article, start article, stub article and list article), however, no ‘A-class’ articles were identified from the list of disease-related Wikidata entries, so this grade was removed from further coding and analysis. Wikipedia article importance is measured on a scale with four levels (top, high, mid and low); WikiProject Medicine’s importance scale typically answers the question, ‘How important is it to Wikipedia’s coverage of this project’s subject area that there should be an article for this topic’.

In the second step, every available, relevant and cited Cochrane review was identified via its associated PubMed ID (PMID). ‘Available’ Cochrane reviews were classified as such on the basis that they were:

1. Published by the Cochrane library and indexed on PubMed.
Relevant’ Cochrane reviews were classified as such on the basis that they were:

1. Published by the Cochrane library and indexed on PubMed.
2. Determined to be relevant to a disease-related Wikipedia article with a corresponding entry on Wikidata.

‘Cited’ Cochrane reviews were classified as such on the basis that they were:

1. Published by the Cochrane library and indexed on PubMed.
2. Determined to be relevant to a disease-related Wikipedia article with a corresponding entry on Wikidata.
3. Cited in that disease-related Wikipedia article.

A Cochrane review was considered to be relevant to a Wikipedia article when the title of the Wikipedia article or its variations in Wikidata appeared in the title or abstract of the review. De-duplication of Cochrane review updates was achieved by applying a title-matching filter to the review list, as updated Cochrane reviews on the same topic retain their titles but are assigned a unique PMID per each update. PMIDs are numerical and sequential; therefore, the title with the largest PMID number will be the most up-to-date version of the review. Instances where an outdated Cochrane review was cited on Wikipedia were recorded (i.e. Cochrane published an ‘updated’ review, but the updated version was not yet cited or reflected in the Wikipedia article). The advantage of this method is that it is computationally efficient, and does not require access to the full texts of the Cochrane review articles such that semantic mapping of full texts can be conducted; the authors envisage that many users may not have access to full-text reviews to enable semantic mapping.

Citation compilation

Once the linking process is complete, the WP2Cochrane system compiles the results into a number of HTML and CSV formatted files which could be easily used by contributors and editors of WikiProject Medicine, without the need of any technical knowledge. The tool creates two different sets of result files. In the first set, all the disease-related Wikipedia articles and their corresponding Cochrane reviews are listed at once. In the second set, the Wikipedia articles and their linkage results are divided between multiple files based on the task force(s) that they belong to. The WikiProject Medicine currently has 17 different task forces (https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine/Task_forces). Members of each task force are assigned a subset of Wikipedia articles most relevant to their speciality to maintain (e.g. cardiology, dermatology, neurology). Our system identifies the task force(s) to which disease-related articles belong and cluster the results accordingly. This feature is designed to further assist Wikipedians to focus their efforts on the articles which are most relevant to their speciality. Figure 2 shows a screenshot of the HTML results page for the neurology task force.

As shown in the figure, results are presented in a seven-column table listing Wikipedia articles and their properties (i.e. title, class, importance, taskforces and corresponding Wikidata disease), and a list of related Cochrane reviews per article. We have used colour coding to indicate the status of each Cochrane review in the list (green: up-to-date and cited, red: up-to-date and not cited, orange: out-of-date and cited, grey: out-of-date and not cited). In cases where a Cochrane review has multiple versions, they are grouped together (showing the same background colour) and listed chronologically (latest version first). The interactive HTML tables are made using the DataTables
(https://datatables.net/) plug-in. This enables users to interactively sort, search and filter the results. All the results files are available online (https://github.com/ajoorabchi/WP2Cochrane).

**Statistical analysis**

Descriptive statistics for: (1) available Cochrane reviews; (2) relevant Cochrane reviews and (3) cited Cochrane reviews were first computed. By aggregating duplicate Cochrane reviews (based on title) and their associated PMIDs, it was possible to identify (4) instances where a disease-related Wikipedia article cited an out of date Cochrane review where a newer one was available (‘out-of-date’ review citations).

We then sought to evaluate the association between outcomes 1–4 and Wikipedia article quality and importance. First, a Spearman’s rank-order correlation was run to assess the relationship between: (a) Wikipedia article quality; (b) Wikipedia article importance and (1) the number of relevant Cochrane reviews; (2) the number of cited Cochrane reviews. Next, our intention was to run a cumulative odds ordinal logistic regression with proportional odds to determine the effect of (1) the number of available Cochrane reviews and (2) the number of cited Cochrane reviews on Wikipedia article quality (including featured, good, B-class, C-class, start, stub and list) and importance (top, high, mid, low). However, preliminary testing revealed a violation of the assumption of proportional odds, as assessed by a full likelihood ratio test comparing the fitted model to a model with varying location parameters. To accommodate this, separate multinomial logistic regression analyses were conducted for each potential expression of the dependent variable (quality and importance), coded in binary format. The independent variables and covariates in the multinomial logistic regression for the different grades of quality were importance, and relevant and cited Cochrane reviews, respectively. The independent variables and covariates in the multinomial logistic regression for the different grades of importance were quality, and relevant and cited Cochrane reviews, respectively.

The p-value for this analysis was set using a Bonferroni-adjusted correction for multiple comparisons at $p < 0.025$ ($0.05/2$ (dependent variables)). All data were analysed using Statistical Analytics Software (Version 18, SPSS Inc., Chicago, IL, USA).
Results

The system of search implementation identified 11,621 disease entries on Wikidata with 4381 corresponding Wikipedia articles.\(^{43}\) Of the 4381 disease-related Wikipedia articles, 1193 cited a Cochrane review, with <1 per cent of these citing an out-of-date review. Of the remaining 3188 (26%) Wikipedia articles, 836 did not cite a Cochrane review, despite at least one ‘relevant’ one being identified.

The characteristics of these 4381 articles with regards to quality and importance are presented in Table 1.

### Characteristics of included articles

In total, 1193 (27%) of the disease-related Wikipedia articles cited one or more Cochrane Reviews. There were 48 instances (1.1% of all citations) where an out of date Cochrane review was cited when a newer version was available (i.e. cited review was ‘out of date’). Thirty-nine of these instances were from separate disease-related Wikipedia articles. There were seven instances where a single disease-related Wikipedia article cited two out of date Cochrane reviews, and one instance where a single disease-related Wikipedia article cited three out of date Cochrane reviews (specifically, the article on ‘Deep Vein Thrombosis’). Due to the low number of citations to out of date Cochrane reviews, this was not included in inferential statistical analyses.

### Spearman analysis

Preliminary analysis showed the relationship to be monotonic between: (a) Wikipedia article quality; (b) Wikipedia article quality and (1) the number of available Cochrane reviews; (2) the number of cited Cochrane reviews, as assessed by visual inspection of a scatterplot. There was a statistically

---

**Table 1.** Characteristics of the included Wikipedia articles, stratified by quality and importance.

| Wikipedia article quality | Featured article | A-class | Good article | B-class | C-class | Start | Stub | List | Unclassified |
|--------------------------|------------------|---------|--------------|---------|---------|-------|------|------|--------------|
|                          | 22               | 0       | 69           | 529     | 944     | 1712  | 906  | 16   | 183          |

| Wikipedia article importance | Top | High | Mid | Low | Unclassified |
|------------------------------|-----|------|-----|-----|---------------|
|                              | 58  | 279  | 2042| 1820| 182           |

**Table 2.** Results of the Spearman analysis evaluating the correlation between Wikipedia article quality and importance, and the number of relevant and cited Cochrane reviews in those articles.

| Spearman’s Rho | Relevant Cochrane reviews | Cited Cochrane reviews |
|----------------|---------------------------|------------------------|
| Wikipedia article quality | -0.362 | -0.328 |
| Wikipedia article importance | -0.436 | -0.314 |
| p < 0.0005 | | p < 0.0005 |
significant, moderate negative correlation between (a) and (b), and (1) and (2), indicating that there were more relevant Cochrane reviews for more important and higher quality Wikipedia articles, and that these articles were also more likely to cite a Cochrane review (Table 2).

**Multinomial logistic regression analyses**

Separate multinomial logistic regression analyses were conducted for each potential expression of the dependent variable. For Wikipedia article quality, this included eight models (featured article vs non-featured article; good article vs non-good article; B-class article vs non-B-class article; C-class article vs non-C-class article; start article vs non-start article; stub article vs non-stub article; list article vs non-list article). For Wikipedia article importance, this included four models (top importance vs non-top importance; high importance vs non-high importance; mid importance vs non-mid importance; low importance vs non-low importance).

Results of the multinomial logistic regression analyses, delineated by article quality and importance, are presented in Tables 3 and 4, respectively.

To summarise the statistically significant findings at the a priori p-value, there was a negative correlation between the number of times that a Wikipedia article classed as being a ‘Good’ article cited a Cochrane review (p = 0.002), and a positive correlation between the number of times that a Wikipedia article classed as a ‘Starting’ article or ‘Stub’ article cited a Cochrane review (p < 0.0005 in both cases). This indicates that higher quality (specifically, ‘Good’) Wikipedia articles were more likely to cite a Cochrane review than lower quality (specifically, ‘Starting’ or ‘Stub’ articles) Wikipedia articles. In contrast, an article’s ‘featured status’ (the highest level on the Wikipedia quality scale) was not explained by its citation of Cochrane reviews, however, this was likely due to the low number of ‘featured’ articles in the dataset (N = 22).

With regard to Wikipedia article importance, there was a statistically significant, negative correlation between the number of relevant Cochrane reviews and Wikipedia articles rated as being of ‘top’ or ‘high’ importance (p = 0.0005 in both cases). There was a statistically significant, positive correlation between the number of relevant Cochrane reviews and Wikipedia articles rated as being of ‘mid’ (p = 0.004) or ‘low’ (p = 0.002) importance. This indicates that more Cochrane reviews are available for more ‘important’ Wikipedia articles. There was also a statistically significant, positive correlation between the number of times that Wikipedia articles of ‘mid’ (p = 0.001) or ‘low’ (p < 0.0005) importance cited a Cochrane review, indicating that less important Wikipedia articles are less likely to cite Cochrane reviews.

**Discussion**

We developed an automated indexing system that links relevant reviews from the Cochrane Library with disease-related Wikipedia articles. This indexing system can be freely and readily accessed and operated online (https://mybinder.org/v2/gh/ajoorabchi/WP2Cochrane/master?urlpath=lab/tree/index.ipynb). The advantage of providing the pipeline in the Jupyter Notebook format is that others can quickly reproduce the results in this article expeditiously, update them if necessary, and potentially conduct alternative analyses of the same dataset. The source code is also hosted on GitHub (https://github.com/ajoorabchi/WP2Cochrane). We encourage users with the necessary technical expertise to add, improve and customise the pipeline as they see fit. The current system can also be leveraged by online information seekers (including healthcare professionals) as they pursue a more comprehensive understanding of a Wikipedia topic using an indexed Cochrane review, and by Wikipedia editors, students or administrators seeking to update or improve disease-related Wikipedia articles.
Table 3. Results of the multinomial logistic regression analyses for article quality, wherein article importance was included as an ordinal-dependent variable and the number of relevant and cited Cochrane reviews were included as continuous covariates.

| Likelihood ratio test (p) | Parameter estimates | 95% confidence interval for Exp(B) |
|---------------------------|---------------------|----------------------------------|
|                           | B       | Sig.  | Exp(B) |                        |
| **Featured article**      |         |       |        |                        |
| 22 of 4174                |         |       |        |                        |
| N Cochrane reviews        | 0.789   |       |        |                        |
| N Cochrane citations      | 0.982   |       |        |                        |
| Importance <0.0005        |         |       |        |                        |
| Nagelkerke R²             | 0.289   |       |        |                        |
| N Cochrane reviews        | –0.002  | 0.787 | 0.998  | 0.986                  |
| N Cochrane citations      | 0.001   | 0.983 | 1.001  | 0.908                  |
| [Importance = Top]        | –22.576 | <0.0005 | 1.569E–10 | 4.34E–11 |
| [Importance = High]       | –20.69  | <0.0005 | 1.034E–09 | 3.39E–10 |
| [Importance = Mid]        | –18.55  |         |        |                        |
| [Importance = Low]        |         |         |        |                        |

| **Good article**          |         |       |        |                        |
| 69 of 4127                |         |       |        |                        |
| N Cochrane reviews        | 0.041   |       |        |                        |
| N Cochrane citations      |         | <0.0005 |       |                        |
| Importance <0.0005        |         |       |        |                        |
| Nagelkerke R²             | 0.256   |       |        |                        |
| N Cochrane reviews        | 0.012   | 0.068 | 1.012  | 0.999                  |
| N Cochrane citations      | –0.161  | 0.002 | 0.851  | 0.767                  |
| [Importance = Top]        | –4.872  | <0.0005 | 0.008  | 0.003                  |
| [Importance = High]       | –2.657  | <0.0005 | 0.07   | 0.028                  |
| [Importance = Mid]        | –1.126  | 0.009 | 0.324  | 0.139                  |
| [Importance = Low]        |         |       |        |                        |

| **B-class**               |         |       |        |                        |
| 529 of 3667               |         |       |        |                        |
| N Cochrane reviews        | 0.412   |       |        |                        |
| N Cochrane citations      | 0.036   |       |        |                        |
| Importance <0.0005        |         |       |        |                        |
| Nagelkerke R²             | 0.157   |       |        |                        |
| N Cochrane reviews        | –0.002  | 0.404 | 0.998  | 0.993                  |
| N Cochrane citations      | –0.063  | 0.048 | 0.939  | 0.882                  |
| [Importance = Top]        | –1.965  | <0.0005 | 0.14  | 0.07                   |
| [Importance = High]       | –2.723  | <0.0005 | 0.066  | 0.047                  |
| [Importance = Mid]        | –0.975  | <0.0005 | 0.377  | 0.297                  |
| [Importance = Low]        |         |       |        |                        |

| **C-class**               |         |       |        |                        |
| 943 of 3253               |         |       |        |                        |
| N Cochrane reviews        | 0.782   |       |        |                        |
| N Cochrane citations      | 0.065   |       |        |                        |
| Importance <0.0005        |         |       |        |                        |
| Nagelkerke R²             | 0.037   |       |        |                        |
| N Cochrane reviews        | –0.001  | 0.781 | 0.999  | 0.995                  |
| N Cochrane citations      | 0.087   | 0.088 | 1.091  | 0.987                  |
| [Importance = Top]        | 1.324   | 0.076 | 3.76   | 0.869                  |
| [Importance = High]       | –0.917  | <0.0005 | 0.4  | 0.295                  |
| [Importance = Mid]        | –0.673  | <0.0005 | 0.51  | 0.435                  |
| [Importance = Low]        |         |       |        |                        |

(Continued)
|                  | Likelihood ratio test (p) | Parameter estimates                                                                 |
|------------------|---------------------------|--------------------------------------------------------------------------------------|
|                  |                           | B          | Sig.       | Exp(B)          | 95% confidence interval for Exp(B) |
|                  |                           | 0         | 0.996      | 1               | 0.996                               |
|                  |                           | 0.613     | <0.0005    | 1.847           | 1.47                                |
|                  |                           | 21.659    | .          | 2549887543      | 2549887543                          |
|                  |                           | 1.762     | <0.0005    | 5.826           | 3.899                                |
|                  |                           | 0.395     | <0.0005    | 1.484           | 1.304                                |
|                  |                           | .          |            | .               |                                     |
|                  |                           | .          |            | .               |                                     |
|                  |                           | .          |            | .               |                                     |
|                  |                           | 0.994     | <0.0005    | 1               | 0.996                               |
|                  |                           | 1.74      | <0.0005    | 5.696           | 2.76                                |
|                  |                           | 20.451    | .          | 762005303       | 762005303                           |
|                  |                           | 4.296     | <0.0005    | 73.429          | 10.252                              |
|                  |                           | 0.459     | <0.0005    | 1.582           | 1.36                                |
|                  |                           | .          |            | .               |                                     |
|                  |                           | .          |            | .               |                                     |
|                  |                           | .          |            | .               |                                     |
|                  |                           | 0.599     | <0.0005    | 0.993           | 0.973                                |
|                  |                           | 0.64      | 0.421      | 1.897           | 0.399                                |
|                  |                           | 15.622    | .          | 6089934.44      | 6089934.44                          |
|                  |                           | -1.697    | 0.072      | 0.183           | 0.029                                |
|                  |                           | -1.229    | 0.06       | 0.293           | 0.081                                |
|                  |                           | .          |            | .               |                                     |
|                  |                           | .          |            | .               |                                     |
Table 4. Results of the multinomial logistic regression analyses for article importance, wherein article quality was included as an ordinal-dependent variable and the number of relevant and cited Cochrane reviews were included as continuous covariates.

|                      | Parameter estimates                                                                 |
|----------------------|-------------------------------------------------------------------------------------|
|                      | B         | Sig. | Exp (B) | 95% confidence interval for Exp(B) |
| Likelihood ratio test (p) |          |      |         |                                   |
| Top importance       | N Cochrane reviews | <0.0005 | 0.11   | 0.597  |
|                      | N Cochrane citations | <0.0005 |         |         |
| Class                | Nagelkerke R² |          |         |         |
|                      | N Cochrane reviews |          |         |         |
|                      | N Cochrane citations |          |         |         |
| [Class=Featured article] |          | <0.0005 | 1.223E–11 | 2.03E–12 |
| [Class=Good article] |          | <0.0005 | 1.492E–11 | 3.059E–12 |
| [Class=C-class]      |          | <0.0005 | 1.346E–10 | 2.971E–11 |
| [Class=Start]        |          | <0.0005 | 2.002E–09 | 2.002E–09 |
| [Class=Stub]         |          | <0.0005 | 0.948    | 0.948   |
| [Class=List]         |          | 0.106  | 1.112   | 1.112   |
| High importance      | N Cochrane reviews | <0.0005 | 0.21    | 0.284  |
|                      | N Cochrane citations | <0.0005 |         |         |
| Class                | Nagelkerke R² |          |         |         |
|                      | N Cochrane reviews |          |         |         |
|                      | N Cochrane citations |          |         |         |
| [Class=Featured article] |          | <0.0005 | 0.983    | 0.978   |
| [Class=Good article] |          | 0.037  | 1.038   | 0.98    |
| [Class=C-class]      |          | <0.0005 | 0.493    | 0.083   |
| [Class=Start]        |          | <0.0005 | 0.721    | 0.143   |
| [Class=Stub]         |          | <0.0005 | 0.444    | 0.099   |
| [Class=List]         |          | 0.335  | 1.397   | 0.311   |
| Mid importance       | N Cochrane reviews | 0.004   | 0.001   | 0.035  |
|                      | N Cochrane citations | <0.0005 |         |         |
| Class                | Nagelkerke R² |          |         |         |
|                      | N Cochrane reviews |          |         |         |
|                      | N Cochrane citations |          |         |         |
| [Class=Featured article] |          | 1.326   | 3.766   | 0.887   |
| [Class=Good article] |          | 0.994  | 2.701   | 0.825   |
| [Class=C-class]      |          | 0.627  | 1.873   | 0.641   |
| [Class=Start]        |          | 0.411  | 1.508   | 0.519   |
| [Class=Stub]         |          | 0.985  | 2.678   | 0.926   |
| [Class=List]         |          | 1.115  | 3.049   | 1.05    |
| Low importance       | N Cochrane reviews | 0.002   | <0.0005 | 0.151  |
|                      | N Cochrane citations | <0.0005 |         |         |
| Class                | Nagelkerke R² |          |         |         |
|                      | N Cochrane reviews |          |         |         |
|                      | N Cochrane citations |          |         |         |
| [Class=Featured article] |          | 19.845  | 415297782 | 415297782 |
| [Class=Good article] |          | <0.0005 | 0.869    | 0.192   |
| [Class=C-class]      |          | <0.0005 | 0.781    | 0.217   |
| [Class=Start]        |          | <0.0005 | 0.492    | 0.138   |
| [Class=Stub]         |          | <0.0005 | 0.218    | 0.062   |
| [Class=List]         |          | <0.0005 | 0.185    | 0.052   |
Using this system, we indexed disease-related Wikipedia articles for which Cochrane reviews were available (‘relevant’ Cochrane Reviews), and evaluated whether those reviews were cited or not cited. We conducted inferential statistical analysis to evaluate the relationship between the quality and importance of a dataset of 4381 disease-related Wikipedia articles, and the number of relevant and cited Cochrane reviews in those articles. Preliminary assumption testing revealed that it was not appropriate to undertake a cumulative odds ordinal logistic regression analysis (due to a violation of the assumption of proportional odds, which likely occurred because the estimated parameters for citation count were not the same for predicting article quality/importance on their respective ordinal scales). As such, separate multinomial logistic regression analyses were conducted. This analysis revealed that there was a negative correlation between Wikipedia articles’ quality and their propensity to cite a Cochrane review. Specifically, articles classed as being ‘Good’ by the Wikipedia community (N = 69) were more likely to cite a Cochrane review (p = 0.002), whereas articles classed as ‘Starting’ (N = 1711) or ‘Stub’ (N = 906) articles by the Wikipedia community were less likely to cite a Cochrane review (p < 0.0005 in both cases). This indicates that higher quality Wikipedia articles (specifically, ‘Good’ articles) were more likely to cite a Cochrane review than lower quality (specifically, ‘Starting’ or ‘Stub’ articles) Wikipedia articles. The relationship between the highest quality Wikipedia article (‘featured’ article status; N = 22) was non-significant, however, this was likely due to the low sample size of this article group. These findings corroborate Wikipedia’s own guidelines for article quality, which specify that higher quality articles must have inline citations from reliable sources. Our findings provide the first objective evidence of the successful implementation of these guidelines in practice.

Assuming that there is a relationship between a disease-related Wikipedia article’s citation of a Cochrane Review and article quality, these findings confirm the validity of Wikipedia’s methods of ensuring article quality in that updating Wikipedia articles that do not currently cite relevant Cochrane reviews is likely to improve their quality. This finding would have implications for the online community of Wikipedia editors as well as for students and professionals in healthcare currently using Wikipedia as an information resource. For example, editors could use the indexing system to expediently identify disease-related Wikipedia articles for which relevant Cochrane reviews are available but not cited, and out of date Cochrane reviews that are currently cited on Wikipedia that need to be updated. Students and professionals who are using Wikipedia to fulfil their healthcare information needs could review the list of non-cited but relevant Cochrane reviews, should they seek to gain further information on a disease-related Wikipedia topic. However, more research is needed to elucidate the relationship between a Wikipedia article’s quality and its citation of a relevant Cochrane review. For the purposes of the present investigation, we relied upon Wikipedia’s own quality grading scale, which is described as a measure of ‘how close we are to a distribution-quality article on a particular topic’. ‘Distribution quality’ as determined by the Wikipedia community may not necessarily reflect scientific quality, as determined by educational or scientific societies; our results should be interpreted with this in mind.

Assuming that there is a relationship between the importance of a disease-related Wikipedia article and the number of available Cochrane reviews, it can be inferred from our findings that any instances of a dissociation between an article deemed by the Wikipedia community to be of ‘top’ or ‘high’ importance and the number of available, relevant Cochrane reviews for that article could potentially direct the Cochrane community to conduct reviews to fills these gaps. For example, the article on ‘Frostbite’ (https://en.wikipedia.org/wiki/Frostbite) is rated as being of ‘high’ importance by the WikiProject Medicine’s dermatology taskforce; the PubMed ‘Clinical Queries’ tool identifies 687 articles relevant to the treatment of ‘Frostbite’ (access date: 1 May 2019), yet no Cochrane reviews exist for this condition. The Cochrane community could systematically review the available scientific literature corpus of this condition to generate a Cochrane review, which could then
be leveraged by the Wikipedia community as a high-quality source of information on this condition. Our analysis revealed that more important Wikipedia articles were more likely to cite a relevant Cochrane review, corroborating this finding.

The dataset of Wikipedia articles included in this analysis were identified through Wikidata. As a machine-readable knowledge base and central storage platform of structured Wikipedia data, we were able to leverage Wikidata to parse metadata about Wikipedia articles in an automated manner. Our dataset of 4381 Wikipedia articles and their associated importance and quality were extracted on the basis that they were classified as a ‘disease’ on Wikidata (The precise Wikidata query is available here: https://w.wiki/3kg). This dataset is relatively smaller than those included in previously conducted studies evaluating Wikipedia articles in the fields of ‘health & fitness’ (18,805 articles, access date: February 2017) and ‘medicine’ (11,314 articles, access date: October 2017), due to the fact that we have undertaken an extra step of filtering to only include those Wikipedia articles which directly describe a disease, and hence was considered relevant in the context of linkage to the Cochrane library. This was deemed appropriate to investigate Cochrane review citations, the research questions for which are population-centred, on the Wikipedia platform.

Of the 4381 disease-related Wikipedia articles, 1193 cited at least one Cochrane review and <1.1 per cent of these cited an out-of-date review. Of practical importance is the finding that a Cochrane review was available for 836 of the 3188 (26%) Wikipedia articles that did not presently cite a Cochrane review. For example, our system identified five relevant Cochrane reviews on ‘mumps’, none of which were cited in the Wikipedia article on this topic. Similarly, our system identified 49 Cochrane reviews on ‘lymphoma’, none of which were cited in the Wikipedia article on this topic. Our analysis would suggest that updating these Wikipedia articles with these Cochrane Reviews would improve their quality on the basis of Wikipedia’s own rating system. In light of the multi-societal disposition to consult the Internet for health-related information and the widely prevalent use of Wikipedia specifically to fulfil these information needs, its value as a free source of knowledge, and the importance of ensuring that this knowledge reflects high-quality scientific evidence, are indisputable. As the most referenced scientific journal among medical Wikipedia articles, the Cochrane Database of Systematic Reviews is an invaluable source of the secondary research cited on Wikipedia. Wikipedia relies heavily on crowd-sourced peer review to ensure the quality of its knowledge corpus. Consequently, facilitating the peer-review process on Wikipedia via the indexing of relevant Cochrane reviews is a public health priority. There have been repeated calls for experienced medical professionals to get more actively involved in improving the accuracy of health-related Wikipedia articles, and Wikipedia has sought to encourage its community to improve the quality of its articles. WikiProject was launched by the Wikimedia foundation for this reason, and WikiProject Medicine, along with its community of editors, might benefit from the indexing system outlined in this article in expediting the editorial process towards improving disease-related Wikipedia articles.

**Limitations**

Despite the novelty of the indexing system, and the insights gained from the analyses outlined in this article, a series of limitations must be acknowledged. First, the indexing system and the associated analyses were limited to the English-language version of Wikipedia. While this is the largest version of Wikipedia, the possibility that other language Wikipedias might have divergent patterns cannot be discounted. Next, and as has already been alluded to, the validity of Wikipedia’s own grading systems of quality and importance may not reflect the sentiments of the scientific community. Further research is required to substantiate whether these scales consistently and reliably reflect scientific robustness (e.g. using the DISCERN criteria) or whether there is a dissociation
of perceived importance between the Wikipedia and scientific or academic communities for different disease-related Wikipedia topics. Finally, because the indexing system described in this report relies on the title of the disease-related Wikipedia article, its terminological derivations and its reroute terms to identify relevant Cochrane reviews, there is an assumption that disease-related Wikipedia articles are both correctly titled (i.e. there are no misspellings) and that these titles reflect the prevailing terminology in scientific research. Due to the potential for incorrect use of search and retrieval ontologies among the scientific community, and the occasional use of non-scientific descriptors as article titles on Wikipedia (e.g. see the Wikipedia article on ‘Charley horse’ to describe a lower limb haematoma; access date: 1 May 2019), it is difficult to establish the precision and recall of the indexing system.

Conclusion

The quality of Wikipedia articles relies on the willingness of experienced and knowledgeable volunteers to devote their time and effort to improve existing Wikipedia articles. Volunteers with WikiProject Medicine help to oversee the medical-related content shared on Wikipedia. Presently, efforts to improve the quality of these articles centre upon appropriate citation of high-quality secondary evidence (including systematic reviews). In this article, we describe an automation system that indexes reviews from the Cochrane Database of Systematic Reviews to relevant disease-related Wikipedia articles. Using this system, we identified an association between the quality of disease-related Wikipedia articles and their citation of Cochrane reviews, whereby higher quality articles were more likely to cite relevant Cochrane reviews. Similarly, a greater number of Cochrane reviews are available for Wikipedia articles deemed to be of greater importance by the community; instances of a dissociation between these variables, whereby no, or fewer reviews are available for more important disease topics, can be used to inform the completion of future reviews. Our indexing system can be accessed via a web browser and is freely available, such that online information seekers can supplement their perusal of disease-related Wikipedia articles with relevant, but uncited, Cochrane reviews where necessary. Similarly, researchers can use this system to replicate, update or further explore our findings, or conduct their own analyses. Technically proficient readers can improve and customise the codebase by forking it on GitHub for their own purposes. Finally, the potential impact of article improvements on Wikipedia should also be considered through the lens of how articles are improved and evidence is shared on the platform. Given that many ‘important’ Wikipedia articles are ‘watched’ by numerous active Wikipedia volunteers, adding one new sentence of evidence and a citation to a Wikipedia article (e.g. evidence from a Cochrane Review) may generate increased Wikipedia editor traffic and result in numerous article improvements by other volunteers, which may result in significant article improvements.

In future, we plan to enhance the usability and impact of our system by:

(a) Development of an NLP-based component capable of measuring the semantic similarity between the textual content of a Cochrane review and a Wikipedia article. This extension will enable us to rank and filter candidate Cochrane reviews according to their semantic relatedness to the given Wikipedia article, and hence reduce the amount of screening work required from the Wikipedians.

(b) Development of a WP2Cochrane robot. There are currently over 2000 software robots running in Wikipedia. These robots carry out a wide range of repetitive tasks helping Wikipedians with the maintenance of millions of articles. Our indexing system currently stores its results in multiple HTML and CSV files. The WP2Cochrane robot, once approved by the WikiProject
Medicine community, will be able to edit the Wikipedia articles’ talk pages directly, and periodically update them with a list of relevant Cochrane reviews. Talk pages (also known as discussion pages) are administration pages where editors can discuss improvements to articles (https://en.wikipedia.org/wiki/Help:Talk_pages).

Author’s note

Jennifer Dawson is also affiliated with CHEO Research Institute, Ottawa, Ontario, Canada.

Acknowledgements

The authors acknowledge Mr Peter Megyesi as the lead developer of a user-facing application that houses the system (the ‘SciScanner’ app).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This research was co-funded by the European Regional Development Fund (ERDF) under Ireland’s European Structural and Investment Funds Programmes 2014-2020 under the ‘SciScanner’ project title, and the Health Research Board (ARPP-A-2018-002).

ORCID iD

Cailbhe Doherty https://orcid.org/0000-0002-5284-856X

References

1. Fenley S. Navigation and visualisation tools usage in large internet and multimedia resources. Berlin: Springer, 2011, pp. 97–104.
2. Gavali MY, Khismatrao DS, Gavali YV, et al. Smartphone, the new learning aid amongst medical students. J Clin Diagn Res 2017; 11(5): JC05–JC08.
3. Lee K, Hoti K, Hughes JD, et al. Dr Google is here to stay but health care professionals are still valued: an analysis of health care consumers’ internet navigation support preferences. J Med Internet Res 2017; 19(6): e210.
4. Sandars J. Technology and the delivery of the curriculum of the future: opportunities and challenges. Med Teach 2012; 34(7): 534–538.
5. Del Fiol G, Workman TE and Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. JAMA Intern Med 2014; 174(5): 710–718.
6. Ahamad A, Wallner P, Salenius S, et al. Information needs expressed during patient-oriented oncology consultations: quantity, variation, and barriers. J Cancer Educ 2019; 34: 488–497.
7. Lim BT, Butow P, Mills J, et al. Information needs of the Chinese community affected by cancer: a systematic review. Psycho-Oncology 2017; 26(10): 1433–1443.
8. Sullivan S. Google now handles at least 2 trillion searches per year, https://searchengineland.com/google-now-handles-2-999-trillion-searches-per-year-250247 (2016, accessed 21 February 2019).
9. Ramaswami P. A remedy for your health-related questions: health info in the knowledge graph. Google, 2015, https://blog.google/products/search/health-info-knowledge-graph
10. Richards L. Why Wikipedia is still visible across Google’s SERPs in 2018, https://www.searchengine-watch.com/2018/11/13/why-wikipedia-is-still-visible-across-googles-serps-in-2018/ (2018, accessed 21 October 2019).
11. Laurent M and Vickers TJ. Seeking health information online: does Wikipedia matter. *J Am Med Inform Assoc* 2009; 16(4): 471–479.

12. Wikipedia. Wikipedia:About, https://en.wikipedia.org/wiki/Wikipedia:About (2019, accessed 13 March 2019).

13. Company AAA. Alexa Top 500, www.alexa.com (2019, accessed 27 February 2019).

14. Heilman JM, Kemmann E, Bonert M, et al. Wikipedia: a key tool for global public health promotion. *J Med Internet Res* 2011; 13(1): e14.

15. Masukume G, Kipersztok L., Das D, et al. Medical journals and Wikipedia: a global health matter. *Lancet Glob Health* 2016; 4(11): e791.

16. Shafee T, Masukume G, Kipersztok L., et al. Evolution of Wikipedia’s medical content: past, present and future. *J Epidemiol Community Health* 2017; 71(11): 1122–1129.

17. Allahwala UK, Nadkarni A and Sebaratnam DF, Wikipedia use amongst medical students – new insights into the digital revolution. *Med Teach* 2013; 35(4): 337.

18. Kingsley K, Galbraith GM, Herring M, et al. Why not just Google it? An assessment of information literacy skills in a biomedical science curriculum. *BMC Med Educ* 2011; 11: 17.

19. Azzam A, Bresler D, Leon A, et al. Why medical schools should embrace Wikipedia: final-year medical student contributions to Wikipedia articles for academic credit at one school. *Acad Med* 2017; 92(2): 194–200.

20. IMS Health. The use of Wikipedia in Health Care, 2014, https://www.iqvia.com.

21. Purdy E, Thoma B, Bednarczyk J, et al. The use of free online educational resources by Canadian emergency medicine residents and program directors. *CIJM* 2015; 17(2): 101–106.

22. Apollonio DE, Broyde K, Azzam A, et al. Pharmacy students can improve access to quality medicines information by editing Wikipedia articles. *BMC Med Educ* 2018; 18(1): 265.

23. Haigh CA. Wikipedia as an evidence source for nursing and healthcare students. *Nurse Educ Today* 2011; 31(2): 135–139.

24. Meek J. Is Wikipedia suitable as a learning resource for nursing and healthcare students? *HIV Nurs* 2016; 16: 65–66.

25. Heilman JM and West AG. Wikipedia and medicine: quantifying readership, editors, and the significance of natural language. *J Med Internet Res* 2015; 17(3): e62.

26. Egle JP, Smeenge DM, Kassem KM, et al. The Internet School of Medicine: use of electronic resources by medical trainees and the reliability of those resources. *J Surg Educ* 2015; 72(2): 316–320.

27. Metcalfe D and Powell J. Should doctors spurn Wikipedia. *JR Soc Med* 2011; 104(12): 488–489.

28. Jemielniak D. *Common knowledge?: An ethnography of Wikipedia.* Stanford, CA: Stanford University Press, 2013.

29. Shachaf P. The paradox of expertise: is the Wikipedia reference desk as good as your library? *J Doc* 2009; 65: 977–996.

30. Frambach JM, Driessen EW, Chan LC, et al. Rethinking the globalisation of problem-based learning: how culture challenges self-directed learning. *Med Educ* 2012; 46(8): 738–747.

31. Callis KL, Christ LR, Resasco J, et al. Improving Wikipedia: educational opportunity and professional responsibility. *Trends Ecol Evol* 2009; 24(4): 177–179.

32. Cochrane. Cochrane and Wikipedia: working together to improve access to health evidence, 2017, https://www.cochrane.org/news/cochrane-and-wikipedia-working-together-improve-access-health-evidence

33. Chapman S. What are Cochrane reviews?, 2014, http://www.evidentlycochrane.net/what-are-cochrane-reviews/

34. Wikipedia. ‘Foundations’ white paper, https://meta.wikimedia.org/wiki/Research:2030 (2019, accessed 13 March 2019).

35. Wikipedia. ‘Knowledge integrity’ white paper, https://meta.wikimedia.org/wiki/Research:2030 (2019, accessed 13 March 2019).

36. Wikipedia. ‘Knowledge gaps’ white paper, https://meta.wikimedia.org/wiki/Research:2030 (2019, accessed 13 March 2019).
37. Jemielniak D, Masukume G and Wilamowski M. The most influential medical journals according to Wikipedia: quantitative analysis. *J Med Internet Res* 2019; 21(1): e11429.

38. Dawson J. Master of science in data science: capstone project submission. University of Virginia Data Science Institute, Charlottesville, VA, 2018, https://ieeexplore.ieee.org/document/8735639

39. Wikipedia. Wikipedia: Assessing articles, https://en.wikipedia.org/wiki/Wikipedia:Assessing_articles (2019, accessed 13 April 2019).

40. Foundation Wikidata. Wikidata, 2019, https://www.wikidata.org/wiki/Wikidata:Main_Page

41. NLM. Cochrane database of systematic reviews included in MEDLINE®. *NLM Tech Bull* 2002; 324: e6.

42. Holtz P, Fetahu B and Kimmerle J. Effects of contributor experience on the quality of health-related Wikipedia articles. *J Med Internet Res* 2018; 20(5): e171.

43. Koppen L, Phillips J and Papageorgiou R. Analysis of reference sources used in drug-related Wikipedia articles. *J Med Libr Assoc* 2015; 103(3): 140–144.

44. Higgins JP and Green S. *Cochrane handbook for systematic reviews of interventions* (Cochrane Book Series). London: Cochrane, 2008, pp. 1–649.

45. Gabarron E, Lau AY and Wynn R. Is there a weekly pattern for health searches on Wikipedia and is the pattern unique to health topics? *J Med Internet Res* 2015; 17(12): e286.

46. Benetoli A, Chen TF, Schaefer M, et al. Professional use of social media by pharmacists: a qualitative study. *J Med Internet Res* 2016; 18(9): e258.

47. Boyd M and Wilson N. Just ask Siri? A pilot study comparing smartphone digital assistants and laptop Google searches for smoking cessation advice. *PLoS ONE* 2018; 13(3): e0194811.

48. Fox S. *Online health search 2006*. Washington, DC: Pew Research Center, 2006.

49. Song H, Omori K, Kim J, et al. Trusting social media as a source of health information: online surveys comparing the United States, Korea, and Hong Kong. *J Med Internet Res* 2016; 18(3): e25.

50. Herbert VG, Frings A, Rehatschek H, et al. Wikipedia – challenges and new horizons in enhancing medical education. *BMC Med Educ* 2015; 15: 32.

51. Brigo F, Lattanzi S, Bragazzi N, et al. Why do people search Wikipedia for information on multiple sclerosis? *Mult Scler Relat Disord* 2018; 20: 210–214.

52. Mathew ME, Joseph A, Heilman JM, et al. Cochrane and Wikipedia: the collaborative potential for a quantum leap in the dissemination and uptake of trusted evidence. *Cochrane Database Syst Rev* 2013; 10: ED000069.

53. Jirschitzka J, Kimmerle J, Halatchliyski I, et al. A productive clash of perspectives? The interplay between articles’ and authors’ perspectives and their impact on Wikipedia edits in a controversial domain. *PLoS ONE* 2017; 12(6): e0178985.

54. Charnock D, Sheperd S, Needham G, et al. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999; 53(2): 105–111.