CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Evolution of wikipedia's medical content: past, present and future
Authors	Shafee, Thomas;Masukume, Gwinyai;Kipersztok, Lisa;Das, Diptanshu;Häggström, Mikael;Heilman, James
Publication date	2017
Original Citation	Shafee, T., Masukume, G., Kipersztok, L., Das, D., Häggström, M. and Heilman, J. (2017) 'Evolution of wikipedia's medical content: past, present and future', Journal of Epidemiology and Community Health, pp. 1-8. doi: 10.1136/jech-2016-208601
Type of publication	Article (peer-reviewed)
Link to publisher's version	http://jech.bmj.com/content/early/2017/08/24/jech-2016-208601 - 10.1136/jech-2016-208601
Rights	© 2017, Article authors or their employer unless otherwise stated in the text of the Article. All rights reserved. No commercial use is permitted unless otherwise expressly granted. This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/ - https://creativecommons.org/licenses/by/4.0/
Download date	2026-10-31 02:09:41
Item downloaded from	https://hdl.handle.net/10468/4893

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com
JECH Online First, published on August 28, 2017 as 10.1136/jech-2016-208601

OPEN ACCESS

Evolution of Wikipedia's medical content: past, present and future

Thomas Shafee,[1] Gwinyai Masukume,[2,3,4] Lisa Kipersztok,[5] Diptanshu Das,[6,7,8,9] Mikael Häggström,[10] James Heilman[11]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jech-2016-208601).

For numbered affiliations see end of article.

Correspondence to
Dr Thomas Shafee, Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia; thomas.shafee@gmail.com, T.Shafee@latrobe.edu.au

Received 2 June 2017
Revised 28 July 2017
Accepted 3 August 2017

ABSTRACT

As one of the most commonly read online sources of medical information, Wikipedia is an influential public health platform. Its medical content, community, collaborations and challenges have been evolving since its creation in 2001, and engagement by the medical community is vital for ensuring its accuracy and completeness. Both the encyclopaedia's internal metrics as well as external assessments of its quality indicate that its articles are highly variable, but improving. Although content can be edited by anyone, medical articles are primarily written by a core group of medical professionals. Diverse collaborative ventures have enhanced medical article quality and reach, and opportunities for partnerships are more available than ever. Nevertheless, Wikipedia's medical content and community still face significant challenges, and a socioecological model is used to structure specific recommendations. We propose that the medical community should prioritise the accuracy of biomedical information in the world's most consulted encyclopaedia.

INTRODUCTION

Why should medical professionals care about Wikipedia?

Wikipedia is one of the most commonly read online sources of medical information, and is consistently among the top 10 most visited websites in the world (currently fifth).[1] As well as being widely read by the general public, it is also used as a source of healthcare information by 50%–70% of physicians[2] and over 90% of medical students.[3] It is additionally used by educators, policymakers and journalists.[4–6] Since the public relies on free online medical information for making health decisions, the accuracy and coverage of Wikipedia's medical information have an immediate real-world impact on public health.[7] The medical community should therefore take responsibility for ensuring its accuracy as an influential health information platform.

Some background on the encyclopaedia and its relatives

Wikipedia is a massive online encyclopaedia with global reach and recognition.[8 9] Its total content has grown rapidly since its inception in 2001, with 44 million articles across 295 languages, including >5.4 million in English as of May 2017 (figure 1A,B). The English-language Wikipedia is the largest and best known project supported by the Wikimedia Foundation (WMF), and will be the main focus of this article. Other projects host open-access images, education materials and structured data.[10 11]

After initial exponential growth of key topics, the English Wikipedia has settled into a slower, linear growth, as more niche topics and current affairs are added. These articles are written and edited by a community of approximately 30 000 editors that make >5 edits per month, and 3000 that make >100 edits per month. This number is down from its peak in 2007, when stricter content guidelines were introduced, but has remained stable over recent years, with a minor increase as easier writing and editing and tools are introduced (figure 1C). The size of the different language versions of Wikipedia is skewed towards English, although less in proportion to the internet as a whole (figure 1D).

Common criticisms of Wikipedia include concerns over content quality, coverage, readability and vandalism. However, much has been done to make Wikipedia's open editing system remarkably robust—from editor culture and policies (eg, increased focus on reliable references)[12 13] to technological improvements (eg, automated software that reverts vandalism).[14] This has been reflected by improvements in perceived accuracy by readers.[15–17] As the encyclopaedia's contents, editors and policies change over time, studies of it can quickly go out of date. This article therefore aims to give an overview of the past, present and possible future of Wikipedia's medical content.

CURRENT STATE OF MEDICAL INFORMATION ON WIKIPEDIA

Content

As of March 2017 there are 30 000 articles on medical topics in English Wikipedia, and another 164 000 in other languages. They are collectively read >10 million times per day.[1] This extreme readership is only approximated by a few other resources, including the US National Institutes of Health and WebMD.[1] Individual medical articles can often have thousands of views per day, although reader numbers for some are strongly affected by news coverage (eg, Zika virus; figure 2A) or seasonal (eg, pneumonia; figure 2B).[18 19]

Wikipedia articles are rated by importance and quality by the communities of editors (online supplementary tables S1 and S2). Top-importance articles include conditions of global significance, such as tuberculosis and pneumonia. High-importance includes common diseases and treatments. Mid-importance encompasses conditions, tests, drugs, anatomy and symptoms. The remaining

To cite: Shafee T, Masukume G, Kipersztok L, *et al*. *J Epidemiol Community Health* Published Online First: [*please include* Day Month Year]. doi:10.1136/jech-2016-208601

Copyright Article author (or their employer) 2017. Produced by BMJ Publishing Group Ltd under licence.

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com

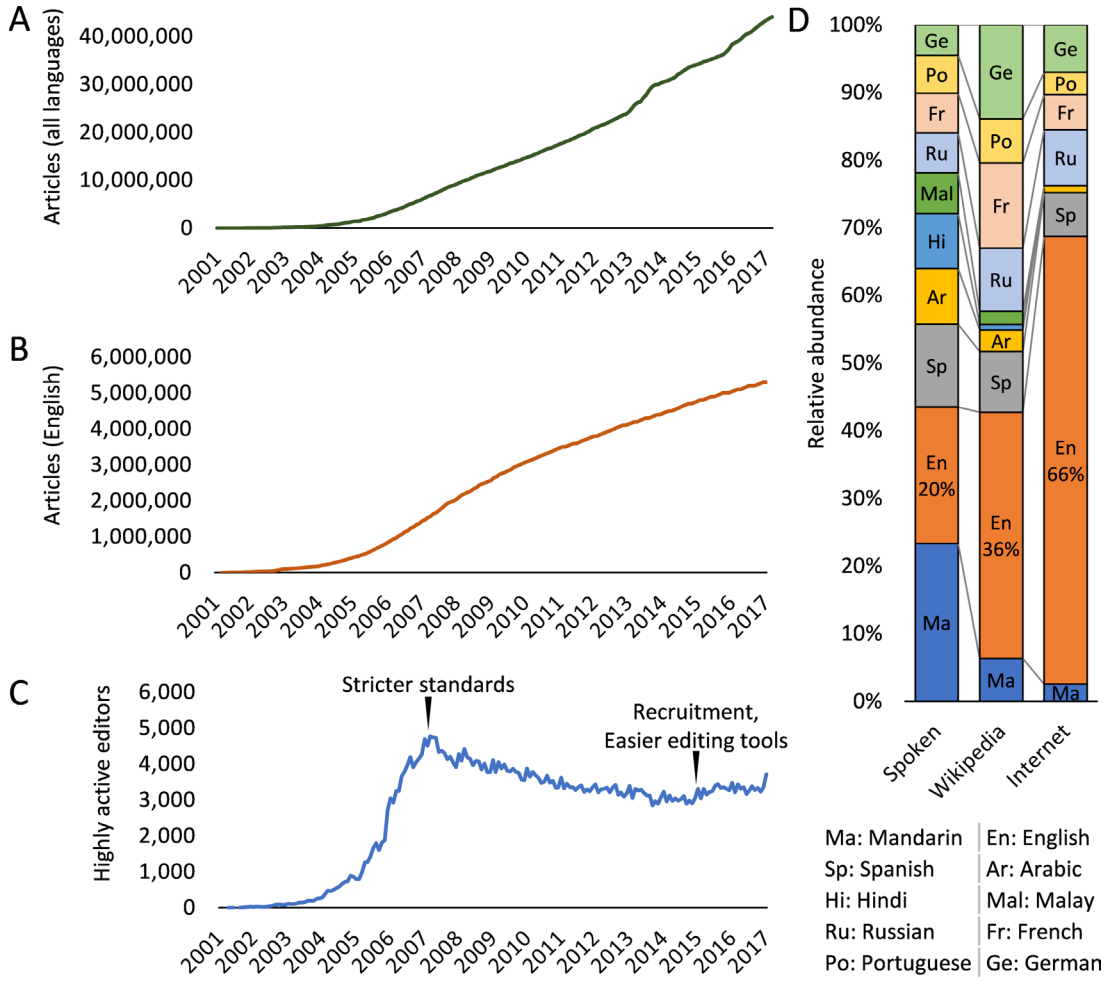


Figure 1 Wikipedia total size and editors. (A) Total number of articles in the encyclopaedia (all languages). (B) Total number of articles in the encyclopaedia (in English). (C) Total number of editors making >100 edits per month (in English). (D) The proportion of the top 10 most spoken languages (native + secondary) compared with the size of different language versions of Wikipedia, and usage on the internet as a whole.

low-importance articles include niche or peripheral medical topics such as laws, physicians and rare conditions. Articles are similarly rated for quality on a scale 'Stub', 'Start', 'C', 'B', 'Good Article' (GA) and 'Featured Article' (FA). The latter two categories are only assigned after an internal peer review process.[20] GAs comprise 0.7% of medical articles and require a single peer reviewer (figure 3A,B). FAs comprise 0.2% of medical articles and have to pass more stringent criteria and often have 5–10 reviewers.

The quality ratings of medical articles are well above Wikipedia's average (figure 3D–F). In particular, 83% of the top-importance medical articles are of 'B-class' quality or above (only 30% Wikipedia-wide) and <1% of the top-importance and high-importance articles remain 'Stub-class' (25% Wikipedia-wide) (figure 3B,E). Over 270 medical articles have been promoted to GA and FA, with around 20 more passing review each year (figure 3C).

External assessment of Wikipedia's overall content quality was found to be comparable to Encyclopaedia Britannica over a decade ago.[21] Comparisons of its medical content with other sources vary for specific subjects, such as pharmacology, psychology or oncology[22–24]; however, some general conclusions can be drawn. Wikipedia's medical content frequently suffers from low readability and errors of omission, despite the fact that included content is relatively high quality and well referenced.[12] Errors are typically not due to deliberate vandalisation or under-qualified editors,[25] but rather that the volunteer editor base is relatively small and so topics are unevenly covered.

Improved referencing for Wikipedia's medical articles has been a strong focus since 2007.[26] Higher quality articles often cite more than a hundred references (figure 4). The majority of references for Wikipedia's medical articles are drawn from reliable sources.[22 27] Furthermore, secondary and tertiary sources (eg, meta-analyses and clinical guidelines) are strongly preferred in order to reflect the accepted medical consensus.[28] Examples from three leading medical journals (*The Lancet*, *New England Journal of Medicine* and *British Medical Journal*) show similar trends, with a high percentage of articles cited by at least one Wikipedia article, and a subset of publications cited multiple times (following a power law).

In general, there is an upward trend in Wikipedia's accuracy and reputation, but completeness and readability are still major limitations.[15–17]

Community

Wikipedia editor communities are organised into approximately 800 currently active 'WikiProjects', which bring together editors interested in a particular topic or process in Wikipedia (online supplementary table S3).[29] WikiProject Medicine was

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com

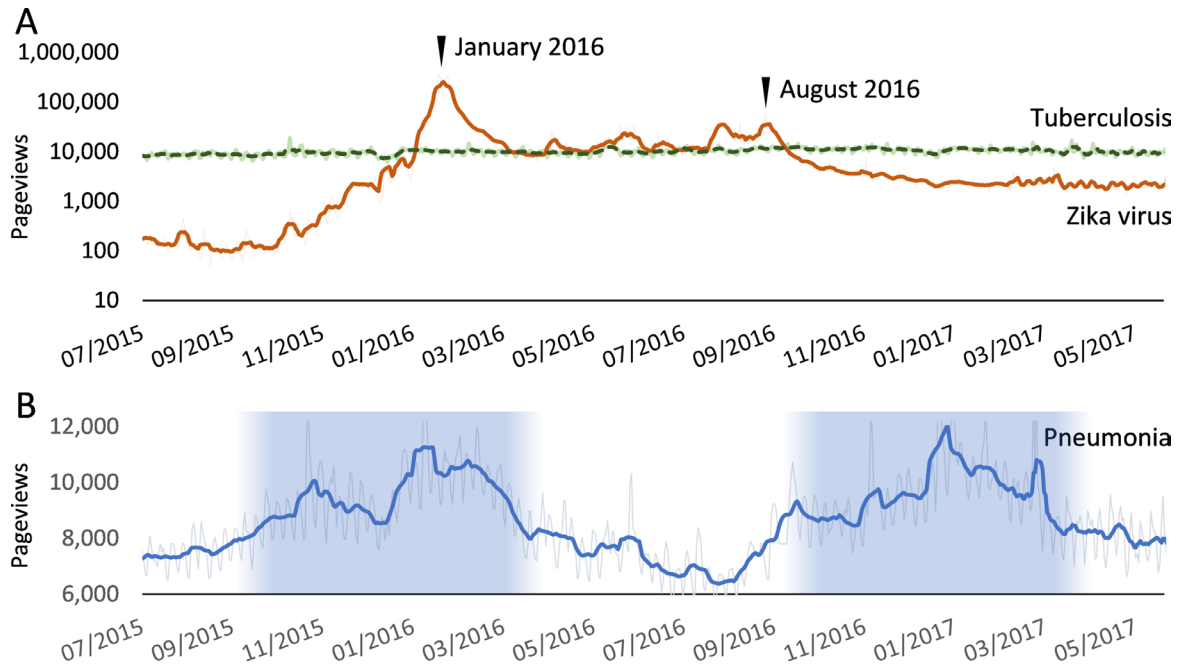


Figure 2 Daily pageviews for medical topics. (A) Most pageviews are relatively stable (eg, tuberculosis), while some are highly dependent on world events (eg, Zika virus). (B) Some seasonal illnesses vary cyclically (eg, pneumonia) by the seasons of the Northern Hemisphere, where approximately 90% of people live. All points smoothed by a 7-day moving average.

one of the first such communities, being founded in 2004 by Jacob de Wolff, MD. It is also one of the most active (consistently in the top 10, online supplementary figure S1) with 130 participants on its discussion forum in any given 90-day window, and a further 700 contributors who edit articles within its scope.[30] The community's overall size has remained relatively constant since 2013. These are largely a mixture of health professionals, researchers and students with an interest in freely available, accurate medical information.[25] Discussion of improvements to content is typically held both at the WikiProject's central discussion page[30] and on the discussion pages for individual articles.[31] The community has since expanded to form the Wiki Project Med Foundation, in 2012, a non-profit corporation working to promote the broader development and distribution of Wikipedia-related medical content.

Collaborations

One of the largest changes in Wikipedia over recent years has been an increasing number of collaborations between the encyclopaedia and the wider biomedical community (figure 5).[32] Dozens of academic publishers, medical institutes and universities have formed temporary and extended partnerships. These benefit the encyclopaedia by providing and improving content, and benefit the partner organisations by increasing impact and awareness as a result of Wikipedia's readership.

One such example is Cochrane, which performs systematic reviews of healthcare research. Since 2012, the organisation has been collaborating with editors to keep articles accurate, up-to-date and evidence-based, and has recruited a 'Wikipedian in residence' to advise and coordinate efforts. For example a 4-month project at the end of 2016 with a team of medical students updated over 100 pages.[33] Cancer Research UK has similarly added hundreds of diagrams and animations.[34] Multiple medical schools have experimented with organising students to edit Wikipedia articles, teaching writing and referencing skills, as well as the value of open access to information.[35]

One of the longest standing biomedical projects has been the 'Gene Wiki' initiative to create Wikipedia articles on all human genes. It began in 2008 with the automated creation of a stub article for every human gene.[36] Initially, 7500 articles were created and 650 updated, with that total since rising above 11 000 as new stubs are automatically added.[36] In 2012 the project formed a further collaboration with the journal *Gene*, whereby articles can be submitted to the journal, then adapted to update the relevant Wikipedia page. This has led to 62 publications integrated into >90 Wikipedia articles.[37] Furthering this precedent, the journal *RNA Biology* requires that new RNA families be added to Wikipedia when published in the journal.[38] A more direct dual-publishing format has been developed by the Public Library of Science (PLOS).[39 40] 'Topic Page' review articles are published in either *PLOS Computational Biology* or *PLOS Genetics*, and then published into Wikipedia to seed a new page, producing 11 such articles since 2014. This model is extended by the *WikiJournal of Medicine*, an open-access medical journal hosted by the WMF, which specialises in this type of integrated publishing.[41] It has additionally put existing Wikipedia articles through academic peer review,[41] following the first experiment in doing so by *Open Medicine* in 2014.[42]

Translation and distribution collaborations increase the impact of content improvement efforts. Translators without Borders, a non-profit organisation, collaborates to translate important medical articles on Wikipedia for the non-English Wikipedias, so far resulting in more than 5.3 million words of translated text in over 100 languages.[43] Several telecom operators in Africa, South East Asia and the Middle East waive data fees for Wikipedia access (Wikipedia Zero), and free mobile apps allow offline storage of Wikipedia medical content for those without reliable internet access.[44] The success of such ventures is highlighted by the 2014 Ebola outbreak. During the early part of the outbreak, teams overhauled the English-language articles on Ebola, and translated them into more than 100 languages. This content was viewed at least 89 million times in 2014 (at a conservative estimate) and was

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com

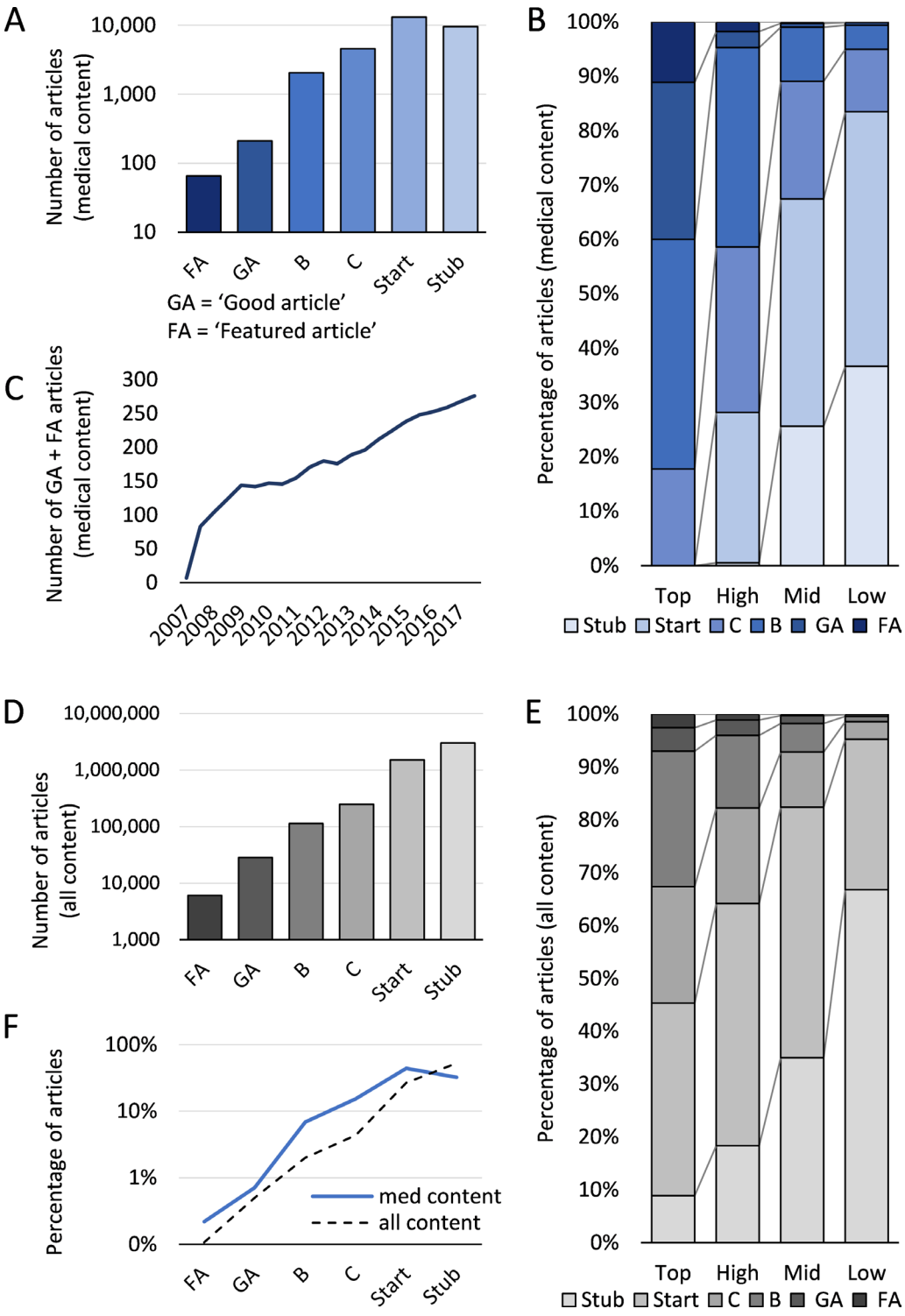


Figure 3 A summary of article quality across all articles in Wikipedia. Articles are assigned quality and importance rankings. 'Featured article' (FA) and 'Good article' (GA) are assigned by internal peer review. (A) Total number of medical articles of each quality ranking. (B) Percentage of medical articles at each quality ranking, separated by article importance (excluding unassessed). (C) The total number of medical articles (FA+GA) over time. (D) Overall number of all articles of each quality ranking Wikipedia-wide. (E) Overall percentage of all articles at each quality ranking, separated by articles importance (excluding unassessed) Wikipedia-wide. (F) A comparison between the overall rankings (dashed line) and those of medical articles (blue line). B, 'B-class article'; C, 'C-class article'.

likely the most used online source for Ebola information in each of the four most affected countries.

Overall, these diverse initiatives add and disseminate high-quality content, as well as introduce Wikipedia editing to people who otherwise would not have contributed.

FUTURE OF MEDICAL CONTENT ON WIKIPEDIA

Challenges

Given Wikipedia's importance as a source of information, it is crucial that it is continually improved and updated. Despite its successes, significant challenges remain to be overcome (table 1).

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com

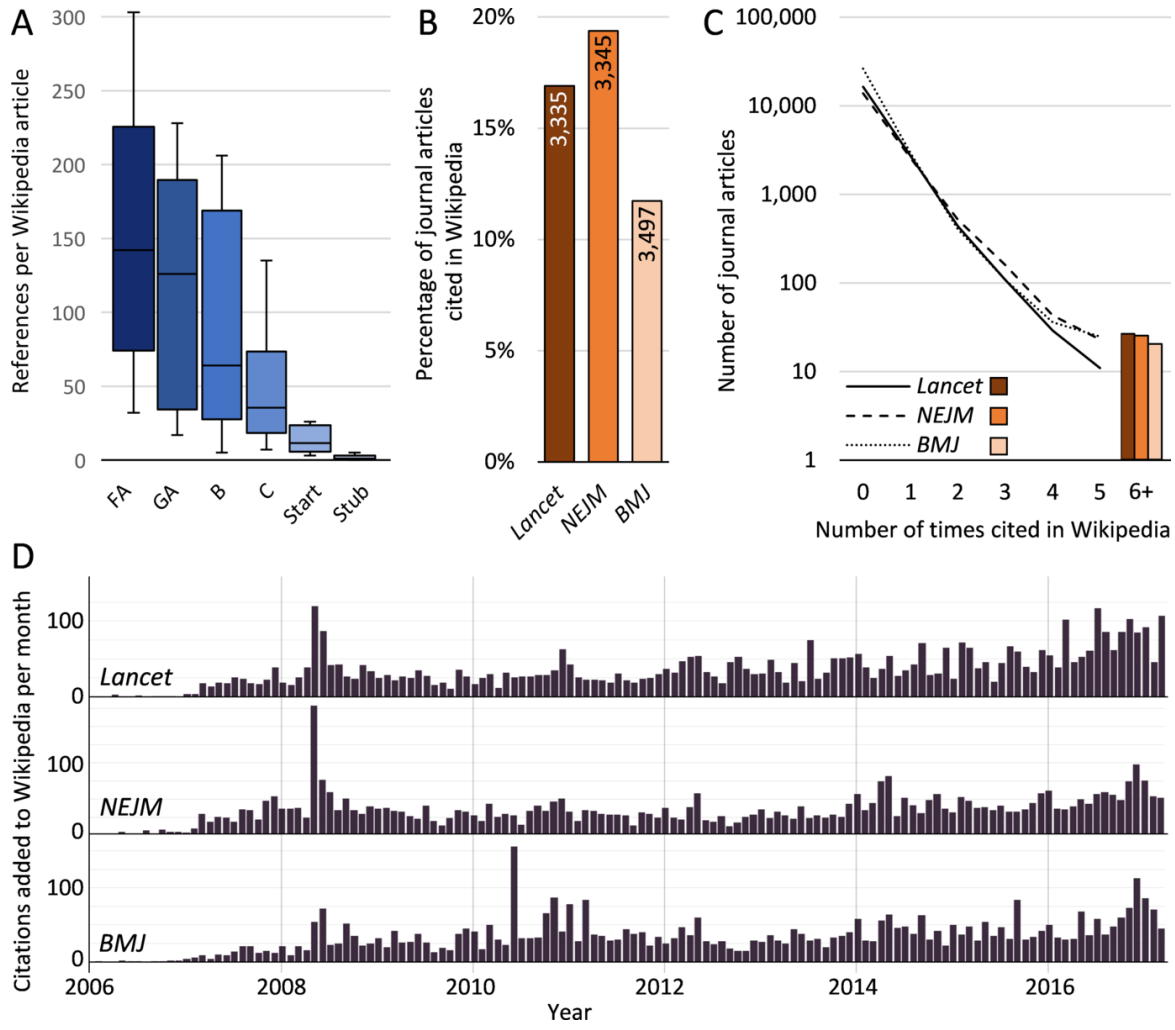


Figure 4 Citation metrics for medical content. (A) Total number of external references per Wikipedia medical article for different article qualities (n=10). FA, 'Featured article'; GA, 'Good article'; B, 'B-class article'; C, 'C-class article'. (B) Absolute numbers and percentage of scholarly articles from three representative high-impact medical journals that have been cited in Wikipedia (up to 13 January 2017). (C) Number of times articles from each representative medical journal have been cited in Wikipedia. (D) Dates when articles are first cited by Wikipedia from selected journals. *Lancet, The Lancet; NEJM, New England Journal of Medicine; BMJ, British Medical Journal*.

Errors and omissions need to be reduced and language often needs to be simplified.[23 45] Content is still heavily skewed towards English. Editor numbers are still insufficient to support expertise on the diversity of specialist topics, and a wider demographic of contributors needs to be recruited.[46] The encyclopaedia norms and bureaucracy need to be simpler and clearer to reduce negative experiences for new editors and to better interface with partner organisations.[47] In general, the rewards for contributing need to justify the time commitment for expert contributors.

The initiatives described in the previous section show how engagement by other organisations can bolster the efforts of Wikipedia's established community of medical content editors. Improved content also generates a positive feedback cycle of increased editing.[48] The WMF is currently developing a strategic plan for the coming 15 years for Wikipedia and its sister projects.[49] In this section, we therefore describe recommendations for established Wikipedia contributors, as well as the medical, research and publishing communities.

Here we use a socioecological framework to make recommendations for the encyclopaedia's main individual, societal, physical and organisational challenges.[50] We suggest an emphasis on collaboration between Wikipedia and external partners, strategies to ensure content quality, better access for low-income and middle-income countries, and improved training and outreach.[49]

Individual

The backbone of Wikipedia is the individual contributions of volunteers. Individual attitudes, behaviours and knowledge consequently have a strong impact.

Wikipedia suffers from many of the same issues in representing the global population as the Science, Technology, Engineering and Mathematics community as a whole.[46] For example, only 10%–20% of editors identify as female.[46] Countering systemic bias towards over-representation of Northern Hemisphere white men will require active recruitment and engagement efforts towards under-represented groups.[51 52] Supporting that recruitment will also require ensuring that the editing interface and culture do not disadvantage under-represented groups.[46] Simultaneously, encouraging engagement by expert contributors is clearly beneficial for ensuring content accuracy.[36 48] Medical and research professionals are busy, and for Wikipedia to be prioritised it is necessary to reiterate Wikipedia's key role as an outreach and public health platform.[44] This includes individually

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com

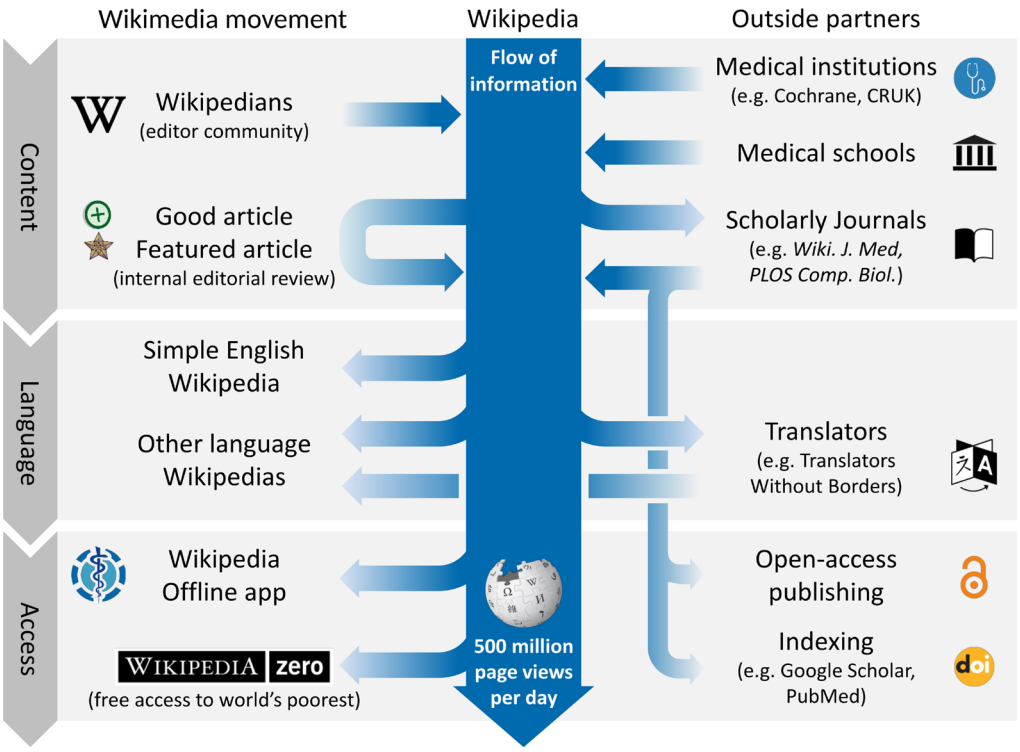


Figure 5 Collaborations between Wikipedia's medical community and outside partners. An outline of collaborations and interactions between the WikiProject Medicine community, institutions, schools, journals and translators (adapted from reference Shafee et al).[32] CRUK, Cancer Research UK; *PLOS Comp Biol*, *PLOS Computation Biology*; *Wiki J Med*, *WikiJournal of Medicine*.

engaging established professionals, as well as training medical and other health sciences students in editing.[34]

The Wikipedia editing community has also had a reputation for being antagonistic and intolerant of mistakes.[53] We strongly endorse the ongoing work by the editor community and WMF to improve community inclusivity.[53 54] It is also worth noting that, when trained to edit Wikipedia, students are remarkably robust to criticism of their errors, highlighting the value of improved tutorials, introductory material and mentoring.[55]

Societal

Wikipedia exists within a broad societal context of multiple overlapping communities, institutions and cultures. Although it is widely used, opinion of Wikipedia in academic circles has often been negative. The consideration of Wikipedia as a low-quality source of evidence discourages contributions and removes the positive reinforcement of recognition for this work.[48] This problem is diminishing as content quality increases and consequently the reputation of Wikipedia improves.[15–17] Rewarding contributors by tracking and recognising their impact would have a large effect on expert contributors.[56]

Tracking and summarising Wikipedia contribution are still difficult. Editors can associate their accounts with their Open Researcher and Contributor ID or Reuters ResearcherID. However, Wikipedia contribution is typically quite different from other forms of authorship. Some Wikipedia articles are written by a small number of authors, like a traditional publication, but it is also common for editors to contribute a small amount to a very large number of articles. New standardised metrics are needed to describe the varied work done in writing, reviewing, improving, debating and illustrating Wikipedia's content, as well as its impact.[56–58]

Medical organisations are already adapting to recognise the diverse ways in which members contribute to public health. The increasingly common Altmetric score includes Wikipedia as one of the indicators of societal impact for academic publications.[59]

Table 1 Summary of some of the main challenges facing Wikipedia and example measures to address them

Challenge	Proposed solution(s)
Content accuracy, readability and language bias[23 45 60] Highly variable article quality Much important information only available to English-language speakers	Greater participation of experts—edits beget edits[48] Content addition through external partnerships Automated language complexity information Tools to recommend articles for translation[60]
Editor numbers, demographic bias and expertise[46 56] Size of the community insufficient to support sufficiently diverse expertise Over-representation of male, white editors from high-income countries Tracking and reward for contribution still underdeveloped	Target diverse editor recruitment Support interface and cultural changes that reduce disadvantage for under-represented groups[46] Develop more sophisticated ways to summarise a contributor's impact[56]
Bureaucracy and policy complexity[47] Wikipedia's norms and policies can be different from those that medical and research professionals are used to, and their descriptions are overly complex.	Simplify and consolidate rules, particularly for new contributors Promote compatible collaboration models with external partners

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com

Recognition and support for Wikipedia content contributions will similarly impact whether experts are able to prioritise engagement in improving Wikipedia.

Physical

Simplifying Wikipedia interfaces can increase article generation by reducing barriers to the technical requirements of contributing to an online encyclopaedia. Reducing complexity and increasing automation of common tasks allow editors to concentrate on content. Additionally, consolidating rules and guidelines will lower barriers for medical practitioners and academics to join or interact with the Wikipedia community. The editing technology, interface and workflow influence who is able to contribute and so influence editor diversity.[46] For example, making it easier for interested editors to find articles that would benefit from translation will help reduce the English-language bias.[60] Automatic feedback on readability of added content can similarly support improved writing. The increasing worldwide internet access, particularly on mobile devices, will further increase use of Wikipedia by academics and medical practitioners less experienced in the technological skills. Further development of tools for non-expert editors will enhance contributions and grow the editor community.

Organisational

The collaborative efforts of medical institutions and Wikipedia are vital to support expert contributors.[44] Wikipedia's relatively extreme egalitarian, open-access and transparent systems can clash with the established norms of medical institutions. Wikipedia's policies and guidelines have become increasingly complex and very different from those that new editors or partner organisations are accustomed to.[47] For example, Wikipedia's protections for anonymity are unusual to a profession where accountability and verified expertise are the norm. Partner organisations have to work out how to interact with Wikipedia systems. Conversely, the Wikipedia community needs to support this by consolidating its policies and guidelines. This will allow new users to avoid accidental errors while they learn the nuances of Wikipedia contribution, and help organisations work out compatible partnerships. Successful models already exist on how to achieve productive collaborations, as discussed in the previous section, and we posit that the public health outcomes are worth the effort.

CONCLUSION

Wikipedia is set to retain its position as a key public health information source. Its content, community, collaboration and challenges have been constantly evolving since it was established in 2001. Proposed socioecological recommendations are most successful when compounded. Many of these issues also involve positive feedback effects; for example, better representation of female editors encourages more to join, and improvement of Wikipedia's reputation encourages expert contribution. Now is a particular period of change as the WMF is currently soliciting feedback to help shape its strategic plan through to 2030. Opportunities for the medical community to shape the encyclopaedia's future stem from individual engagement with its 'anyone can edit' model, and increasing partnerships with the wider medical ecosystem. The medical community must work together to ensure that medical content is accurate in the world's most consulted encyclopaedia.

Author affiliations
[1]Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, Melbourne, Australia
[2]Department of Obstetrics and Gynaecology, University College Cork, The Irish Centre for Fetal and Neonatal Translational Research, Cork, Ireland
[3]Gravida: National Centre for Growth and Development, University of Auckland, Auckland, New Zealand
[4]Division of Epidemiology and Biostatistics, University of the Witwatersrand, School of Public Health, Johannesburg, South Africa
[5]Department of Family Medicine, Oregon Health & Science University, Portland, Oregon, USA
[6]Department of Paediatric Neurology, Universita degli Studi di Roma Tor Vergata, Roma, Italy
[7]Department of Paediatrics, Kothari Medical Centre and Research Institute, Kolkata, India
[8]Department of Pediatrics, ICARE Institute of Medical Sciences and Research, Haldia, India
[9]Working Group, Open Access India, India
[10]Department of Radiology, NU Hospital Group, Trollhättan, Sweden
[11]Department of Emergency Medicine, University of British Columbia, Vancouver, Canada

Contributors TS gathered and analysed data for the figures. GM contributed to analysis for figure 4. All authors (TS, GM, LK, DD, MH, JH) discussed the material, wrote the manuscript, and approved the final version.

Funding The work is not externally funded. The article processing fee was covered by a rapid grant from the Wikimedia Foundation. The Foundation had no involvement in the study design, the collection, analysis and interpretation of the data, writing the report, or in the decision to submit the manuscript for publication.

Competing interests All authors have contributed to Wikipedia articles. TS, DD, MH and JH are current participants in WP:MED. JH is a former and current member of the Wikimedia Foundation board of trustees. All authors are on the editorial board of WikiJournal of Medicine. TS is on the editorial board of PLOS Genetics. The authors do not receive financial compensation for their contributions to these projects.

Provenance and peer review Commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1 Heilman JM, West AG. Wikipedia and medicine: quantifying readership, editors, and the significance of natural language. *J Med Internet Res* 2015;17:e62.
2 Hughes B, Joshi I, Lemonde H, *et al*. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. *Int J Med Inform* 2009;78:645–55.
3 Allahwala UK, Nadkarni A, Sebaratnam DF. Wikipedia use amongst medical students - new insights into the digital revolution. *Med Teach* 2013;35:337.
4 Messner M, South J. Legitimizing Wikipedia: How US national newspapers frame and use the online encyclopedia in their coverage. *Journalism Practice* 2011;5:145–60 .

What is already known on this subject

- Wikipedia is one of the most used medical information resources globally, with immediate public health implications.
- Its model of user-generated content presents unique challenges and opportunities.
- Its content quality is variable but improving, and in need of further expert input.

What this study adds

- We summarise the major trends in how Wikipedia's medical content, community and collaborations have changed since its inception in 2001.
- We raise specific proposals for both the Wikipedia community and medical institutions to help improve the encyclopaedia.

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com

5 Purcell K, Heaps A, Buchanan J, *et al*. How teachers are using technology at home and in their classrooms. *PEW research* 2013 http://www.pewinternet.org/2013/02/28/how-teachers-are-using-technology-at-home-and-in-their-classrooms/ (accessed 8 Apr 2017).
6 Peoples L. The Citation of Wikipedia in Judicial Opinions. *Yale Journal of Law and Technology* 2009;12 1 http://heinonline.org/HOL/LandingPage?handle=hein.journals/yjolt12&div=3&id=&page=
7 Fox S, Duggan M, *Health Online 2013*. PEW research, 2013. http://www.pewinternet.org/2013/01/15/health-online-2013/
8 Masukume G, Kipersztok L, Das D, *et al*. Medical journals and Wikipedia: a global health matter. *Lancet Glob Health* 2016;4:e791.
9 Parve S, Ershadi A, Karimov A, *et al*. Access, attitudes and training in information technologies and evidence-based medicine among medical students at University of Zimbabwe College of Health Sciences. *Afr Health Sci* 2016;16:860–5.
10 Pfundner A, Schönberg T, Horn J, *et al*. Utilizing the Wikidata system to improve the quality of medical content in Wikipedia in diverse languages: a pilot study. *J Med Internet Res* 2015;17:e110.
11 Singh S. Use of wikiversity and role play to increase student engagement during student-led physiology seminars. *Adv Physiol Educ* 2013;37:106–7.
12 Mesgari M, Okoli C, Mehdi M, *et al*. 'The sum of all human knowledge': A systematic review of scholarly research on the content of Wikipedia. *J Assn Inf Sci Tec* 2015;66:219–45 .
13 Priedhorsky R, Chen J, Lam S. (tony) K, et al. Creating, destroying, and restoring value in wikipedia. *Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP' 07*. New York: ACM Press, 2007:259 .
14 Adler BT, de Alfaro L, Mola-Velasco SM, *et al*; Wikipedia Vandalism Detection: Combining Natural Language, Metadata, and Reputation Features. Gelbukh A, ed. *Computational Linguistics and Intelligent Text Processing. Springer Berlin Heidelberg*, 2011:277–88 .
15 Soules A. Faculty perception of Wikipedia in the California State University System. *New Library World* 2015;116:213–26.
16 Aibar E, Lladós-Masllorens J, Meseguer-Artola A, *et al*. Wikipedia at university: what faculty think and do about it. *The Electronic Library* 2015;33:668–83.
17 Jemielniak D, Aibar E. Bridging the gap between wikipedia and academia. *J Assoc Inf Sci Technol* 2016;67:1773–6.
18 Generous N, Fairchild G, Deshpande A, *et al*. Global disease monitoring and forecasting with Wikipedia. *PLoS Comput Biol* 2014;10:e1003892.
19 McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. *PLoS Comput Biol* 2014;10:e1003581.
20 Viégas FB, Wattenberg M, McKeon MM. The Hidden Order of Wikipedia. In. *Lecture Notes in Computer Science*:445–54 .
21 Giles J. Internet encyclopaedias go head to head. *Nature* 2005;438:900–1.
22 Reavley NJ, Mackinnon AJ, Morgan AJ, *et al*. Quality of information sources about mental disorders: a comparison of Wikipedia with centrally controlled web and printed sources. *Psychol Med* 2012;42:1753–62.
23 Kräenbring J, Monzon Penza T, Gutmann J, *et al*. Accuracy and completeness of drug information in Wikipedia: a comparison with standard textbooks of pharmacology. *PLoS One* 2014;9:e106930.
24 Leithner A, Maurer-Ertl W, Glehr M, *et al*. Wikipedia and osteosarcoma: a trustworthy patients' information? *J Am Med Inform Assoc* 2010;17:373–4.
25 Farič N, Potts HW. Motivations for contributing to health-related articles on Wikipedia: an interview study. *J Med Internet Res* 2014;16:e260.
26 Rasberry L. Wikipedia: what it is and why it matters for healthcare. *BMJ* 2014;348:g2478.
27 Haigh CA. Wikipedia as an evidence source for nursing and healthcare students. *Nurse Educ Today* 2011;31:135–9.
28 Wright JG. A practical guide to assigning levels of evidence. *J Bone Joint Surg Am* 2007;89:1128–30.
29 Tinati R, Luczak-Roesch M, Shadbolt N, *et al*; Using WikiProjects to Measure the Health of Wikipedia. *Proceedings of the 24th International Conference on World Wide Web - WWW' 15 Companion*. New York, USA: ACM Press , 2015:369–70 .
30 James R. WikiProject Medicine: Creating Credibility in Consumer Health. *J Hosp Librariansh* 2016;16:344–51.
31 Viegas F, Wattenberg M, Kriss J, *et al*. Talk Before You Type: Coordination in Wikipedia. *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*: IEEE:78 .
32 Shafee T, Heilman J, Masukume G, *et al*. Wikipedia's medical content: a new era of collaboration. *Wikimedia Blog* 2016 https://blog.wikimedia.org/2016/10/29/wikipedia-medical-content/ (accessed 11 Jan 2017).
33 de Haan S. Cochrane Community. Wikipedia: an important dissemination tool for Cochrane. 2017 http://community.cochrane.org/news/wikipedia-important-dissemination-tool-cochrane (accessed 19 May 2017).
34 Byrne J. WikiProject CRUK Final report. Cancer Research UK, 2015. https://en.wikipedia.org/wiki/Wikipedia:WikiProject_CRUK/Final_report_2014-15(accessed 19 May 2017).
35 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;92:194–200.
36 Burgstaller-Muehlbacher S, Waagmeester A, Mitraka E, *et al*. Wikidata as a semantic framework for the Gene Wiki initiative. *Database* 2016;2016:baw015.
37 Tsueng G, Good BM, Ping P, *et al*. Gene Wiki Reviews-Raising the quality and accessibility of information about the human genome. *Gene* 2016;592:235–8.
38 Butler D. Publish in Wikipedia or perish. *Nature* 2008;6.
39 Wodak SJ, Mietchen D, Collings AM, *et al*. Topic pages: PLoS Computational Biology meets Wikipedia. *PLoS Comput Biol* 2012;8:e1002446.
40 Luk A. Continuing to bridge the journal-wikipedia gap: introducing topic pages for PLOS genetics. *PLOS Biologue* 2017 http://blogs.plos.org/biologue/2017/04/12/continuing-to-bridge-the-journal-wikipedia-gap-introducing-topic-pages-for-plos-genetics/ (accessed 17 May 2017).
41 Shafee T, Das D, Masukume G, *et al*. WikiJournal of Medicine, the first Wikipedia-integrated academic journal. *WikiJournal of Medicine* 2017;4.
42 Maskalyk J. Modern medicine comes online: How putting Wikipedia articles through a medical journal's traditional process can put free, reliable information into as many hands as possible. *Open Med* 2014;8:e116–9.
43 Tran M. Translators fight the fatal effects of the language gap. The Guardian, 2012. http://www.theguardian.com/global-development/poverty-matters/2012/apr/11/volunteers-translation-language-health-messages(accessed 22 May 2017).
44 Wikipedia Zero. Wikimedia Foundation. https://wikimediafoundation.org/w/index.php?title=Wikipedia_Zero&oldid=107138 (accessed 7 Jul 2017).
45 Brigo F, Erro R. The readability of the English Wikipedia article on Parkinson's disease. *Neurol Sci* 2015;36:1045–6.
46 Shane-Simpson C, Gillespie-Lynch K. Examining potential mechanisms underlying the Wikipedia gender gap through a collaborative editing task. *Comput Human Behav* 2017;66:312–28.
47 Heaberlin B, DeDeo S. The Evolution of Wikipedia's Norm Network. *Future Internet* 2016;8:14.
48 Aaltonen A, Seiler S. Cumulative Growth in User-Generated Content Production: Evidence from Wikipedia. *Manage Sci* 2016;62:2054–69.
49 Wikimedia Strategy development (Cycle 2). meta.wikimedia.org https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Cycle_2 (accessed 11 May 2017).
50 Fielding JE, Teutsch S, Breslow L. A Framework for Public Health in the United States. *Public Health Rev* 2010;32:174–89.
51 Keystones to foster inclusive Knowledge Societies. Unesco 4 March, 2015. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/internet_draft_study.pdf(accessed 17 May 2017).
52 Strugnell J. Profiles: Kudos for female Antarctic researchers. *Nature* 2016;536:148.
53 Morgan JT, Bouterse S, Walls H, *et al*. Tea and sympathy: crafting positive new user experiences on wikipedia. *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW' 13*. New York, USA: ACM Press, 2013:839 .
54 Narayan S, Orlowitz J, Morgan JT, *et al*. Effects of a Wikipedia Orientation Game on New User Edits. *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing-CSCW'15 Companion*, 2015 .
55 Shane-Simpson C, Che E, Brooks PJ. Giving Psychology Away: Implementation of Wikipedia Editing in an Introductory Human Development Course. *Psychology Learning & Teaching* 2016;15:268–93.
56 Allen L, Scott J, Brand A, *et al*. Publishing: Credit where credit is due. *Nature* 2014;508:312–313.
57 Adler BT, de Alfaro L, Pye I, *et al*. Measuring author contributions to the Wikipedia. *Proceedings of the 4th International Symposium on Wikis-WikiSym' 08*. New York, USA: ACM Press, 2008:1 .
58 Arazy O, Stroulia E, Ruecker S, *et al*. Recognizing contributions in wikis: Authorship categories, algorithms, and visualizations. *Journal of the American Society for Information Science and Technology* 2010;21:n/a.
59 Sugimoto CR, Work S, Larivière V, *et al*. Scholarly use of social media and altmetrics: A review of the literature. *J Assoc Inf Sci Technol* 2017;15:379.
60 Wulczyn E, West R, Zia L, *et al*. Growing Wikipedia Across Languages via Recommendation. *Proc Int World Wide Web Conf* 2016;2016:975–85 .

Downloaded from http://jech.bmj.com/ on October 10, 2017 - Published by group.bmj.com



Evolution of Wikipedia's medical content: past, present and future

Thomas Shafee, Gwinyai Masukume, Lisa Kipersztok, Diptanshu Das, Mikael Häggström and James Heilman

J Epidemiol Community Health published online August 28, 2017

Updated information and services can be found at:
http://jech.bmj.com/content/early/2017/08/24/jech-2016-208601

These include:

References This article cites 39 articles, 3 of which you can access for free at:
http://jech.bmj.com/content/early/2017/08/24/jech-2016-208601#BIBL

Open Access This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See:
http://creativecommons.org/licenses/by/4.0/

Email alerting service Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections Articles on similar topics can be found in the following collections

Open access (308)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/