Letter to Editor

Asian Pacific Journal of Tropical Medicine

Quality of monkeypox information in Wikipedia across multiple languages

Eric David B. Ornos[1], Jared Gabriel L. Dela Rosa[1], Jea Giezl N. Solidum[1], Jervy P. Garcia[2], Erika P. Ong[1], Rafael Lorenzo G Valenzuela[1], Ourlad Alzeus G. Tantengco[1✉]

[1]College of Medicine, University of the Philippines Manila, Ermita, Manila, Philippines

[2]Department of Family and Community Medicine, Eastern Visayas Medical Center, Tacloban City, Philippines

In recent years, the Internet has become the primary source of health information for the general population, which may be attributed to improvements in digital technology and Internet accessibility[1]. Since the World Health Organization declared the coronavirus disease (COVID-19) as a pandemic in March 2020[2], digital information has gained more importance, as seen through the rapid growth in the number of people searching online[3]. As seen from previous infectious disease outbreaks, the recent increase in monkeypox cases might compel individuals worldwide to broaden their searches for relevant virtual health information[3,4].

Interestingly, digital health information seekers prefer contents published in Wikipedia since it usually yields the top of search results, exceeding other well-reputed online sites such as the Mayo Clinic and the World Health Organization[5]. However, Wikipedia's content has long been subjected to questionable quality, and many researchers continue to doubt the platform[6]. Due to the potential threat of monkeypox to human health, it is essential to assess the quality of online health information available to the public. This study used a highly utilized and validated information tool to determine the quality and reliability of Wikipedia articles on monkeypox written in the 30 most spoken languages in the world.

One of the authors searched Wikipedia for articles on monkeypox written in the 30 most spoken languages globally, according to the 2022 Ethnologue 200 list[7]. The search was conducted on June 23, 9:00 am, and was saved in pdf. format for analysis to avoid any possible discrepancies resulting from the constant updates and changes among Wikipedia entries.

The authors evaluated the quality and reliability of the monkeypox articles in Wikipedia using the DISCERN instrument, a 16-item assessment tool developed by the British Library and the United Kingdom National Health Service Executive Research & Development Programme to assess healthcare-related websites and online resources[8]. Descriptive statistics and Spearman's rank-order correlation analysis were performed using GraphPad Prism version 9 (GraphPad Software, San Diego, CA).

Wikipedia had entries for monkeypox disease in 19 of the 30 most spoken languages (Table 1). These languages are used by 5 571 million speakers worldwide. The top languages without Wikipedia entries for monkeypox were (in millions): Hindi (n=602), Urdu (n=231), Nigerian Pidgin (n=120), Marathi (n=99), Yue Chinese (n=86), Tagalog (n=82), Wu Chinese (n=81), Hausa (n=77), Swahili (n=71), Javanese (n=68), and Western Punjabi (n=66).

Table 1 shows the results of the DISCERN instrument quality assessment of monkeypox Wikipedia entries. The mean total DISCERN instrument score was (35.1±6.2), (23.5±3.5) for the first section, (9.4±2.5) for the second section, and (2.2±0.7) for the last section. The mean number of words for the Wikipedia pages was (2 471±2 205), while the mean number of references was (25.9±23.1).

Seven Wikipedia pages had an overall quality of 3 (fair quality with useful source of information but some limitations). These pages were for the monkeypox entries written in English, French, Bengali, Indonesian, German, Tamil, and Korean. On the other hand, three Wikipedia pages (Turkish, Vietnamese, and Egyptian) received an overall quality of 1 (low quality with extensive shortcomings). The rest of the Wikipedia entries received a score of 2 (poor quality with shortcomings) for overall quality. Some of the issues identified in the Wikipedia articles were a limited description of disease symptoms and transmission, little or no discussion on prevention and treatment options, and unclear or incomplete information for references.

The monkeypox Wikipedia pages with the highest number of references were entries from English (n=107), German (n=50), French (n=44), and Russian (n=40) languages. Six Wikipedia pages provided readers with more than five external links (English, French, Tamil, Egyptian Arabic, Mandarin Chinese, and Japanese).

We found that DISCERN score correlated significantly with the

Table 1. Summary of DISCERN scores and characteristics of Wikipedia entries for monkeypox from 19 of the 30 most spoken languages.

Language	Total speakers (million)	DISCERN instrument				Number of words	Number of references	Number of external links
		Sec I	Sec II	Sec III	Total			
English	1452.0	29	12	3	44	9557	107	7
Mandarin Chinese	1118.0	25	8	2	35	3549	30	5
Spanish	548.3	22	9	2	33	1422	15	1
French	274.1	26	9	3	38	6165	44	7
Standard Arabic	274.0	23	8	2	33	2021	19	0
Bengali	272.7	24	13	3	40	2238	20	0
Russian	258.2	24	8	2	34	3846	40	0
Portuguese	257.7	26	9	2	37	2632	32	2
Indonesian	199.0	25	13	3	41	1297	20	4
Standard German	134.6	29	12	3	44	4395	50	0
Japanese	125.4	19	6	2	27	1392	10	5
Telugu	95.7	22	8	2	32	1235	14	0
Turkish	88.1	19	6	1	26	492	6	0
Tamil	86.4	24	11	3	38	976	10	6
Vietnamese	85.3	23	6	1	30	574	8	1
Korean	81.7	27	13	3	43	1556	18	1
Iranian Persian	77.4	23	12	2	37	1210	16	1
Egyptian Arabic	74.8	13	6	1	20	526	2	6
Italian	67.9	24	9	2	35	1883	31	4

Note: The other 11 in the 30 most spoken languages have no Wikipedia entry for monkeypox.

number of words (r=0.603, P=0.006) and the number of references (r=0.712, P<0.001) of the Wikipedia articles. Similarly, the total number of speakers correlated significantly with the number of words (r=0.689, P=0.001) and the number of references (r=0.547, P=0.015). DISCERN score did not correlate with the total number of speakers.

This study assessed the content quality of Wikipedia entries for monkeypox virus across multiple languages. We found that (1) Wikipedia pages on monkeypox were poor in quality and lacked essential details, (2) 11 out of 30 most spoken languages had no Wikipedia entries on monkeypox, (3) most pages had references and external links provided, and (4) the number of words and references were positively correlated with quality of the Wikipedia page based on the DISCERN tool.

There is a need to improve the quality of health information about monkeypox in Wikipedia. These content variations reflect socioeconomic, health, and research disparities between high and low- to middle-income countries. Health professionals and organizations must address inequities in access to accurate health information to help improve global health outcomes.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Funding

Authors' contributions

EDBO and OAGT conceptualized the study. EDBO, JGLDR, JGNS, JPG, EPO, RLGV, and OAGT performed the investigation, data curation, formal analysis, and writing the original draft. EDBO and OAGT reviewed and edited the manuscript. All authors approved the publication of this manuscript.

References

[1] van der Vaart R, Drossaert C. Development of the digital health literacy instrument: Measuring a broad spectrum of health 1.0 and health 2.0 skills. *J Med Internet Res* 2017; **19**(1): e27. doi: 10.2196/jmir.6709.

[2] Afshar ZM, Ebrahimpour S, Javanian M, Koppolu V, Vasigala VK, Hasanpour AH, Babazadeh A. Coronavirus disease 2019 (COVID-19), MERS and SARS: Similarity and difference. *J Acute Dis* 2020; **9**: 194-199.

[3] Effenberger M, Kronbichler A, Shin JI, Mayer G, Tilg H, Perco P. Association of the COVID-19 pandemic with internet search volumes: A Google TrendsTM analysis. *Intern J Infect Dis* 2020; **95**: 192-197.

[4] Tausczik Y, Faasse K, Pennebaker J, Petrie P. Public anxiety and information seeking following the H1N1 outbreak: Blogs, newspaper articles, and Wikipedia visits. *Health Commun* 2011; **27**(2): 179-185.

[5] Chrzanowski J, Sołek J, Fendler W, Jemielniak D. Assessing public interest based on Wikipedia's most visited medical articles during the SARS-CoV-2 outbreak: Search trends analysis. *J Med Intern Res* 2021; **23**(4): e26331. doi: 10.2196/26331.

[6] Jemielniak D. Wikipedia: Why is the common knowledge resource still neglected by academics? *Gigascience* 2019; **8**(12): giz139. doi: 10.1093/gigascience/giz139.

[7] Ethnologue. *What are the top 200 most spoken languages?* [Online]. Available from: https://www.ethnologue.com/guides/ethnologue200. [Accessed on 7 July 2022].

[8] Charnock D, Shepperd S, Needham G, Gann R. DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999; **53**(2): 105-111.

Publisher's note

The Publisher of the *Journal* remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.