

Chapter 7

Are Wikipedia Articles Reliable Learning Resources in Problem-Based Learning Curricula?

Samy A. Azer

7.1 Introduction

Since its creation, Wikipedia has evolved as an important resource for patients, the general public, students and health professionals. Wikipedia is the largest online encyclopaedia with 4,541,520 articles in English (English Wikipedia, 2014). Important features of Wikipedia may include: (1) it is available free of charge for everyone without prior registration or membership; (2) it covers different aspects of knowledge and entertainment including arts, biography, geography, history, mathematics, science, society, technology, business and health; (3) it enables users to add their contributions and therefore these tools get enriched as more people use them; and (4) it has been seen by the general public, students, and health professionals as important sources for information.

With the changes introduced to medical curricula such as the introduction of problem-based learning (PBL) and the accommodation of self-regulated learning as part of the curriculum design, it has been noted that most medical students tend to search easily accessible online resources such as Google, and Wikipedia websites for their 'learning issues' (Alegría, Boscardin, Poncelet, Mayfield, & Wamsley, 2014; Patil et al., 2014; Petty, 2013). While the new changes in the curriculum aimed at enhancing students' skills to critically search for knowledge from several resources rather than study the content of a particular textbook, the development of technology and its availability have shifted students' search for knowledge from paper-based resources to online resources. Although several online resources have been created for medical and health professionals such as Medscape (eMedicine),

S.A. Azer (✉)

Medical Education Department, Curriculum Development and Research Unit,
College of Medicine, King Saud University, Riyadh, Saudi Arabia
e-mail: azer2000@optusnet.com.au

UpToDate Inc., the Merck Manual Medical Library and PubMed (Azer, 2014; Tyson, 2000), there is evidence that medical students usually prefer to start by searching general online resources such as Google and Wikipedia to find answers for their queries (Kingsley et al., 2011). There is also evidence that physicians use the Internet far more than the general public (Masters, 2008). The work of Hughes, Joshi, Lemonde, and Wareham (2009) showed that 53 % of Internet visits were made by junior physicians and were directed mainly to Google and Wikipedia. Despite awareness of participants about information credibility risks related to these online resources and the risk of poor quality information obtained, the junior physicians in the study preferred easily accessed resources such as those provided by Wikipedia. They viewed Wikipedia and Google as important sources of medical information. These views have also been found to apply to medical and allied health students (Guarino et al., 2014; Kolski, Arlt, Birk, & Heuwieser, 2013; Prasannan, Gabbur, & Haughton, 2014).

However, these online resources may not be created by academics or qualified experts. Furthermore, both Wikipedia, and YouTube do not appoint expert editors to assess work submitted and materials are published online without prior peer-review or expert evaluation. The absence of prior review in assessing the quality and the scientific content of Wikipedia articles raises several questions regarding the adequacy and scientific accuracy of these resources. Two recent studies showed that physicians using online resources may become less vigilant towards potential errors even when it contradicts their existing knowledge (Lau & Coiera, 2008; Westbrook, Gosling, & Coiera, 2005). Recently Schmidt et al. (2014) demonstrated that media information about a disease gained from a source such as Wikipedia can cause internal medicine residents to misdiagnose similar-looking clinical cases. The problem of availability bias may arise from exposure to media-provided information about a disease, causing diagnostic errors (Schmidt et al., 2014). The bias's effect is apparently associated with nonanalytic reasoning and can be counteracted by reflection. By this, we mean the tendency of residents to make shortcuts and jump into conclusions without careful analysis, possibly due to the effect of fast knowledge obtained from media information.

Despite their wide use by medical students and junior doctors, there are limited studies exploring the accuracy of these learning resources and whether they have attained the standards required in scholarly/academic resources. To determine the suitability of Wikipedia articles for medical students as a source for information, it is important to determine whether articles are scientifically accurate, up-to-date, free from errors and whether there are no gaps or deficiencies in the information provided. Also, it is important to determine their suitability for medical students and whether they are written at a reading level appropriate for college students rather than for lay persons in the general public. Therefore, the purpose of this study was to evaluate the suitability of the nervous system articles available on the English-language Wikipedia that might be used by medical students as part of their learning resources. To answer this question Wikipedia articles were assessed with respect to: (1) scientific accuracy and comprehension; (2) frequency of updating and quality of references; (3) reliability; and (4) readability.

7.2 Methods

This study analysed the nervous system articles on the English Wikipedia database (<http://en.wikipedia.org>).

7.2.1 *Wikipedia Articles*

7.2.1.1 Study Design

To identify the topics on the nervous system and its disorders, five medical textbooks (Table 7.1) and eMedicine (Medscape) website (www.emedicine.com) were searched. These books were used because they are recommended by most medical schools, have been reviewed in peer-reviewed journals such as the British Medical Journal and the New England Medical Journal, have been written and edited by medical experts, are regularly updated every 3–4 years, and several editions have been produced over the past 30–60 years. eMedicine (Medscape) is a professional educational website written and regularly updated by medical consultants and eminent clinicians.

The aims of searching these resources were to: (1) ensure that topics needed by medical students in their undergraduate course in relation to the nervous system have been identified and included in the search; and (2) use these resources as standardized reference in the assessment of the accuracy and quality of information provided in Wikipedia articles. The chapters on the nervous system were revised and key topics were identified by three evaluators (the author, a medical consultant and professor of medical education, plus two medical graduates). The lists of identified key topics were discussed in a meeting. The three evaluators agreed upon a final list covering 42 topics.

Table 7.1 Medical textbooks used as a standardized reference in evaluating Wikipedia articles

Andreoli TE, Benjamin IJ, Griggs RC, Ewing EJ. Andreoli and Carpenter's Cecil Essentials of Medicine. 8th Edition, Philadelphia: Saunders, 2010
Colledge NR, Walker BR, Ralston SH. Davidson's Principles & Practice of Medicine, 21st Edition, Edinburgh: Elsevier, Churchill Livingstone, 2010
Kumar P, Clark M. Kumar and Clark's Clinical Medicine. 8th Edition. Edinburgh: Elsevier, 2013
Longo DJ, Fauci AS, Kasper DL, Hauser SL, Jameson JL, Loscalzo J. Harrison's Principles of Internal Medicine. 18th Edition. New York: McGraw-Hill, 2012
McPhee SJ, Papadakis MA, Rabow MW. Current Medical Diagnosis & Treatment. 50th Anniversary Edition, New York: McGraw-Hill, Lange, 2011

7.2.1.2 Searching Wikipedia

The Wikipedia website (<http://en.wikipedia.org/wiki/Wikipedia>) was searched on 20 May 2014 for the 42 topics. Topics were printed out and a photocopy was given to each evaluator. The aim of using a photocopied version of the articles rather than the electronic version was to ensure that all evaluators were using the same version. This is particularly important as Wikipedia articles undergo continuous changes.

7.2.1.3 Instrument Used in Assessing Accuracy

The rating instrument used in this study was a modified version of the DISCERN instrument (Azer, 2014, see Appendix C for the modified version used). The DISCERN project was funded by the British Library and the NHS Executive Research and Development Program from 1996 to 1997 (DISCERN Project, 1999). The instrument consists of 15 questions, plus a question about the overall evaluation of the document, and was designed for the evaluation of the different aspects of healthcare related websites and information about treatment options. For example, the DISCERN instrument was used in assessing online resources on epidural anesthesia (Jaffe, Tonick, & Angell, 2014), mental health (Grohol, Slimowicz, & Granda, 2014), colorectal cancer (Grewal & Alagaratnam, 2013), and inflammatory bowel disease (Van der Marel et al., 2009).

However, the original DISCERN instrument is not suitable for evaluating Wikipedia articles as it was not designed to assess scientific accuracy in given information, the inclusion of illustrations, figures, tables or multimedia to support the topics or whether there are gaps or deficiencies in the information given. These deficiencies drove the need to modify the DISCERN instrument as discussed in an earlier publication (Azer, 2014). The modified DISCERN instrument is comprised of ten questions and aims at providing a comprehensive assessment of Wikipedia articles in regard to: (1) aims of the article and the adequacy of subtitles used; (2) scientific accuracy of information provided and if there were any personal views; (3) degree of balancing of different parts, and whether sources of information were provided; (4) regular updating the article and if there were gaps or deficiencies in the article that need to be completed; (5) images, figures and tables provided; and (6) the overall rating of the article.

The original DISCERN scoring system has been used. Each question is rated on a 5-point scale, where 1 corresponds to 'no', 3 corresponds to 'partially yes', and 5 corresponds to 'yes'. For the last question, 1 corresponds to 'serious or extensive shortcomings', 3 corresponds to 'potentially important but not serious shortcomings', and 5 corresponds to 'minimal shortcomings'.

7.2.1.4 Piloting the Study

Prior to applying the instrument to the Wikipedia articles, the use of the modified DISCERN instrument was piloted with the aim to: (1) orient the evaluators to the different items of the instrument and the scoring system; (2) ensure that evaluators

were able to use the instrument; and (3) enhance the evaluators skills in applying the instrument through feedback on their assessment. For piloting purpose, ten Wikipedia articles other than those included in the study were selected and each evaluator was asked to evaluate them using the instrument independently. Articles that were scored differently were discussed and a resolution was reached.

Ten additional Wikipedia articles were selected and were again evaluated by the three evaluators using the modified DISCERN instrument as described earlier. The agreement between the evaluators was in the range of 75–85 %, which was considered satisfactory.

7.2.1.5 Assessing Wikipedia Articles on the Nervous System

The 42 articles identified were evaluated independently by the three evaluators using the modified DISCERN instrument. Interrater agreement between the evaluators for each item in the modified DISCERN instrument was calculated using Cohen's kappa score (Kharbanda et al., 2012; Tsivgoulis et al., 2013).

7.2.1.6 Assessing References

The list of references at the end of each article was evaluated by the three evaluators independently. The aims of evaluating the references were to assess if the authors used appropriate resources to construct each article and what type of references they used. Therefore, the following points were considered in the evaluation: total number of references; number of peer-reviewed journals; number of educational guidelines and proceedings from professional societies, textbooks, professional and general websites; and others (such as news and media). Articles written for academic purposes need to rely on up-to-date, peer-reviewed references such as scientific/medical articles, educational guidelines and proceedings produced by professional societies rather than cite general references such as general websites, non-peer-reviewed articles, magazine articles, and news.

7.2.1.7 Frequency of Wikipedia Article Updates

Resources written for academic purposes are expected to be regularly reviewed and updated. Such reviews usually aim at enhancing the quality of content, adding up-to-date information and recent developments as well as related references. The frequency of updating articles was assessed through the 'view history' button next to 'search' at the top right part of each article. Information collected included: (1) date created; (2) total number of revisions; (3) total number of authors; (4) average time between edits; (5) average edits per month; (6) revisions in the last 12 months; and (7) average edits per user.

7.2.1.8 Assessing Readability

The aim of assessing the readability was to evaluate whether the reading level was appropriate for medical students. Two methods were used in assessing readability: Flesch-Kincaid grade level and Coleman-Liau index (Vargas, Chuang, Ganor, & Lee, 2014). The score of readability is an indicator of the number of years of education that a person needs to be able to understand the text on the first reading. For example, a score of ten indicates that a tenth grade student can easily understand the topic. An online calculator, provided by Readability Formulas (<http://www.readabilityformulas.com/free-readability-formula-tests.php>), was used. Based on the instructions given, a random sample of 150–60 words were copied from the beginning, middle and the end of each article and placed into the space provided by the programme. Headings, external links, images and numbers of citations were omitted from the text used prior to conducting the calculation. The reading scores for each part were recorded and the mean and standard deviation were calculated for each article.

7.2.2 Statistical Analysis

The mean, standard deviation, minimum and maximum were calculated. To assess the degree to which different evaluators agreed in their assessment, the Cohen's kappa interrater reliability was calculated. Correlation between the DISCERN scores and the number of updates and number of peer-reviewed references were also calculated. The aim was to assess whether the number of updates, the number authors/reviewers and references were related to the improvement of the article quality or not.

7.3 Results

7.3.1 Depth and Accuracy of Articles

Table 7.2 summarizes the number of pages, the scores calculated using the modified DISCERN instrument and the number of images, illustrations, tables and media/ audios of each Wikipedia article. The number of pages ranged from one page for the article on smell to 36 pages for the article on stroke indicating that topics varied in regard to details given and depth of discussion. Also, this may be because some articles were incomplete, had gaps or deficiencies in their content and needed further work. Considering the number of pages of articles and the frequency of updating/reviewing since first created, there is evidence that some topics were of less interest to Wikipedians and were less frequently reviewed or improved compared to other articles.

Table 7.2 Accuracy score, number of images, illustrations, tables, readability scores and number of references of each Wikipedia nervous system article included in the study

Topic title	Number of pages	Accuracy score (mean±SD)	Number of images/illustrations	Number of audios/media/videos	Number of tables	Readability Flesch-kincaid grade level (mean±SD)	Readability Coleman-Liau Index (mean±SD)	Total references	External links (EL)/Further reading (FR)
Amyotrophic lateral sclerosis	20	33.67±1.15	1 Illustration 2 MRI scans	1	2	13.80±3.40	13.00±2.00	88	2 EL
Ataxia	12	26.33±1.52	0	0	0	16.90±1.60	15.00±2.00	37	17 EL 2 FR
Autonomic nervous system	12	23.33±1.15	1 Illustration	0	11	17.90±1.65	14.33±1.15	20	5 EL
Basal ganglia	14	29.00±1.00	5 Illustrations 2 Pathology	1	1	15.56±2.15	12.33±2.08	24	4 EL
Brainstem	7	29.00±1.00	9 Pathology	1 Media related	0	12.40±0.55	12.67±0.57	4	7 EL
Brainstem glioma	7	22.00±1.00	0	0	0	12.23±1.81	13.67±1.15	28	6 EL 5 FR
Broca's area	9	18.67±0.57	3 Illustrations	1 Media related	1	15.90±2.94	13.33±2.51	24	4 EL
Cerebellum	27	32.67±1.52	18 Illustrations	1 Media related	0	15.10±2.49	12.00±1.73	69	2 EL 3 FR
Cerebral circulation	4	24.33±0.57	3 Illustrations	0	0	15.03±2.40	14.67±2.51	0	2 EL
Cerebral hemisphere	6	22.00±1.00	4 Illustrations	1 Media related	1	13.86±1.58	15.67±0.57	8	0
Cerebrospinal fluid	11	26.00±1.00	1 Photo 5 Illustrations	2 Media related	0	11.80±2.10	11.33±1.15	25	3 EL
Coma	11	26.67±0.57	1 Photos	2 Videos	1	15.26±1.45	11.67±2.08	26	0 EL
Cranial nerves	20	30.33±1.52	27 Illustrations 1 Pathology	1 Media related	1	13.90±1.64	10.00±1.00	33	5 EL

(continued)

Table 7.2 (continued)

Topic title	Number of pages	Accuracy score (mean±SD)	Number of images/illustrations	Number of audios/media/videos	Number of tables	Total	Readability Flesch-kincaid grade level (mean±SD)	Readability Coleman-Liau Index (mean±SD)	Total references	External links (EL)/Further reading (FR)	
Dementia	26	28.67±0.57	2 Illustrations	0	1	4	11.46±2.05	11.33±2.51	79	1 EL	
			1 Map distribution of disability-aged life year for Alzheimer								
Drizziness	4	24.67±4.72	0	0	0	0	15.16±5.68	15.67±4.72	11	2 EL	
Duchenne muscular dystrophy	16	35.33±1.52	1 Illustration	1 Media related	0	3	12.63±0.80	12.33±0.57	45	5 EL	
			1 Histopathology								
Encephalitis	5	21.00±1.00	1 MRI Scan	0	0	1	16.40±5.99	17.33±2.30	7	4 EL	
Encephalopathy	4	15.00±1.00	0	0	0	0	17.30±2.45	17.00±2.64	3	1 EL	
											1 FR
Epilepsy	26	33.33±1.52	3 Photos	1 Media related	0	8	11.66±1.96	11.67±1.15	123	1 EL	
			2 Illustrations	1 Video							2 FR
			1 Map-epidemiology								
Ganglion	3	16.67±0.57	1 Histopathology	0	0	2	14.83±0.47	14.00±1.00	6	0 EL	
			1 Photo								
Hearing	7	23.67±1.52	2 Illustrations	0	0	2	11.7±1.45	13.33±1.52	15	3 EL	
											2 FR
Lower motor neuron lesion	2	17.33±0.57	1 Illustration	0	0	1	16.63±3.35	17.00±2.00	1	1 EL	
Meningitis	22	27.00±1.00	2 Illustrations	0	1	7	15.06±1.74	12.67±1.15	71	2 EL	
			2 Photos								
			1 Microbiology								
			1 Histopathology								

Motor neuron	5	22.00 ± 1.00	1 Histopathology	0	1	2	17.00 ± 5.04	14.00 ± 2.00	5	2 FR	
Multiple sclerosis	28	38.00 ± 1.00	2 MRI Scans	0	0	13	13.90 ± 1.65	12.00 ± 1.73	114	2 EL	
			1 Histopathology								1 FR
			7 Illustrations								
			2 Photos								
			1 Map-epidemiology								
Myasthenia gravis	14	30.00 ± 1.00	2 Photos	0	0	9	15.96 ± 2.75	14.67 ± 1.15	30	4 EL	
			5 Illustrations								
			1 Photo								
			1 CT-Scan								
Myelin	6	29.00 ± 1.00	2 Illustrations	0	0	3	12.06 ± 1.27	13.33 ± 1.15	13	9 EL	
			1 Photo								
Myopathy	5	24.67 ± 0.57	0	0	0	0	11.00 ± 1.63	12.33 ± 0.57	5	2 EL	
Neuromuscular junction	10	29.67 ± 1.52	1 Photo	0	0	3	15.10 ± 1.82	15.33 ± 2.08	16	1 EL	
			2 Illustrations								3 FR
Neuron	19	31.67 ± 1.52	8 Illustrations	1 Media related	0	12	14.13 ± 2.00	11.33 ± 0.57	26	10 EL	
			3 Histopathology								6 FR
Olfaction	13	24.67 ± 0.57	2 Photos	1 Media related	0	5	14.73 ± 0.90	13.67 ± 0.57	41	14 EL	
			2 Illustrations								1 FR
Parkinson's disease	30	29.67 ± 1.52	4 Illustrations	0	0	14	15.93 ± 1.85	14.67 ± 1.15	111	4 EL	
			7 Photos								
			2 MRI Scan								
			1 Map-epidemiology								
Peripheral neuropathy	10	26.33 ± 0.57	1 Photo	0	0	1	16.63 ± 4.30	15.67 ± 3.05	25	2 EL	
Sensory neuron	5	15.67 ± 0.57	0	0	0	0	15.60 ± 2.35	13.33 ± 1.52	11	0 EL	

(continued)

Table 7.2 (continued)

Topic title	Number of pages	Accuracy score (mean±SD)	Number of images/illustrations	Number of audios/media/videos	Number of tables	Readability Flesch-kincaid grade level (mean±SD)	Readability Coleman-Liau Index (mean±SD)	Total references	External links (EL)/Further reading (FR)
Smell	1	10.33±0.57	0	0	0	No enough text to test	No enough text to test	0	0 EL
Spinal cord	13	23.67±0.57	16 Illustrations 4 Photos	1 Media related	0	11.53±1.35	10.67±0.57	2	7 EL
Spinal cord injury	20	31.00±1.00	3 Illustrations 2 Photos	0	0	14.56±1.38	12.33±1.15	71	4 EL
Stroke	36	36.00±1.00	1 Photo 1 Map-epidemiology 1 ECG 3 CT-Scan 1 Pathology 2 Illustrations 2 Histopathology	1 Media related	0	13.66±2.11	11.67±0.57	158	1 EL 2 FR
Taste	21	23.67±0.57	1 Illustration	1 Media related	0	10.96±1.30	11.33±0.57	87	3 EL 21 FR
Upper motor neuron lesion	2	22.00±1.00	1 Illustration	0	0	14.13±1.85	12.67±0.57	3	2 EL
Vertigo	9	27.33±1.15	0	1 Media related	0	15.06±4.45	12.00±2.00	38	0 EL
Wernicke's area	5	22.67±0.57	1 Illustration	1 Media related	0	16.00±1.50	13.67±0.57	14	0 EL

The minimum DISCERN score was 10.33 ± 0.57 (mean \pm SD) for the article titled ‘smell’ and the maximum score was 38.00 ± 1.00 for the article titled ‘multiple sclerosis’. The mean score for the 42 articles was 25.88 ± 5.97 . To summarize the scores, there were 10 articles scoring 30 or higher, 26 articles scoring 20–29, and 6 articles 10–19 (the maximum score was 50). Top scored articles such as the article on ‘multiple sclerosis’ was covered on 28 pages, had 13 images, illustrations and photos, 114 references, 2 external links and one further reading. Also the article on ‘stroke’ scored 36.00 ± 1.00 , was covered on 36 pages, had 12 images, illustrations, photos and media related, 158 references, 1 external link and 2 further readings.

On the other hand, articles with the lowest scores such as the article on ‘smell’ scoring 10.35 ± 0.57 was one page only, had no images, illustrations, tables, photos or multimedia. Also the article had no references or external links. The article on ‘encephalopathy’ is another example, scoring 15.00 ± 1.00 , had no images, illustrations, photos or tables, only three references, one external link and one further reading. It was not possible to measure the readability for articles comprised of one page only.

Although the articles followed the template of Wikipedia for medical/health related articles, some articles were incomplete and most articles were deficient in the areas of: (1) disease pathogenesis; (2) clinical picture; and (3) management of nervous system diseases. Agreement between evaluators was calculated by Cohen’s kappa interrater correlation; the range for the mean \pm SD for the scores were 0.65 ± 0.10 to 0.79 ± 0.12 .

7.3.2 Article References

Table 7.2 summarizes the total number of references, external links and further readings for each Wikipedia topic. The total number of references for the 42 articles was 1517 and the number varied from 0 to 158 references; 36.12 ± 38.82 (mean \pm SD). There was weak correlation between the DISCERN scores and the number of total references. This suggests that the absolute number of references was not a good measure for assessing the quality of an article. This is particularly important as not all references were peer-reviewed articles and educational guidelines produced by professional bodies were lacking in the list of references in most articles.

Common problems found in citations and the list of references can be summarized as follows: (1) citation of wrong references, and failing to cite the appropriate references; (2) incomplete references (for example, missing journal or book title, missing year, volume or page numbers); (3) inconsistencies in the way the references are written; (4) failure to include guidelines of professional societies/associations; and (5) several statements in articles are missing appropriate references as in-text citations.

7.3.3 *Frequency of Revisions*

Table 7.3 summarizes key information about articles, their history in regard to date created, number of revisions, number of authors, average time between edits, edits per month and in the last 12 months. It is obvious from the article's history that the date of creation varied. For example, while the earlier article 'stroke' was created on the 16th of April 2001, the most recent article 'brainstem glioma' was created on the 25th of January 2008.

While there was moderate correlation between the DISCERN score and the total number of revisions ($R^2=0.38$) and the total number of authors ($R^2=0.42$), there was weak correlation between the DISCERN score and the average edits per month ($R^2=0.10$).

7.3.4 *Article Readability*

To calculate readability, two methods were used. Table 7.2 shows the readability scores (mean \pm SD) for each article as calculated by Flesch-Kincaid grade level and Readability Coleman-Liau index. The range of readability using the first method was in the range of 10.96 ± 1.30 to 17.90 ± 1.65 , while the second method showed a range of 10.00 ± 1.00 to 17.33 ± 2.30 . The article on 'smell' did not have enough text to calculate readability. A good correlation was found between the scores calculated by the two methods, $R^2=0.650$. The mean score for all articles was 14.21 ± 2.91 on using the first method and 13.02 ± 2.74 for the second method. These scores of readability indicate that Wikipedia articles were geared to college level.

7.4 Discussion

The aim of this study was to evaluate the quality and accuracy of content of Wikipedia articles as learning resources commonly used by medical students. To evaluate these resources, we specifically evaluated the accuracy, clarity, quality of information and readability of Wikipedia articles on the nervous system. A total of 42 Wikipedia articles covering the nervous system diseases were evaluated using the modified DISCERN instrument (Azer, 2014).

Although Wikipedia articles followed the template created by Wikipedia for medical/health articles, most articles were deficient in addressing disease pathogenesis, clinical picture, and management of nervous system diseases. The accuracy of articles as measured by the modified DISCERN instrument had a mean score of 25.88 ± 5.97 ; only ten articles scored 30 or higher and over 50 % of articles scored 20–29 out of 50. Although images, illustrations, photos, and multimedia were incorporated in some articles to enhance the educational value of articles, the quality of

Table 7.3 Wikipedia articles on the nervous system: article history, date created, latest edit, number of revisions, number of authors, and frequency of revisions

Number	Topic title	Page size (Bytes)	Date created	Total number of revisions	Total number of distinct authors	Average time between edits (days)	Average edits per month	Revision in last 12 months	Average edits per user
1	Amyotrophic lateral sclerosis	75,414	28 January 2003	3267	1678	1.3	24	339	1.9
2	Ataxia	35,361	3 June 2001	666	419	7.1	4.3	41	1.6
3	Autonomic nervous system	31,031	6 January 2003	984	620	4.2	7.2	68	1.6
4	Basal ganglia	40,320	4 October 2002	829	366	5.1	5.9	76	2.3
5	Brainstem	15,309	25 May 2003	725	409	5.5	5.5	103	1.8
6	Brainstem glioma	20,187	25 January 2008	87	52	26.2	1.2	10	1.7
7	Broca's area	25,025	21 February 2002	428	265	10.3	3	29	1.6
8	Cerebellum	75,828	25 April 2002	1738	760	2.5	11.9	131	2.3
9	Cerebral circulation	6291	18 July 2004	51	34	70.5	1	2	1.5
10	Cerebral hemisphere	8864	18 December 2003	333	218	11.4	2.7	32	1.5
11	Cerebrospinal fluid	26,331	2 January 2002	589	377	7.7	3.9	54	1.6
12	Coma	34,236	2 June 2001	1156	814	4.1	7.4	93	1.4
13	Cranial nerves	47,486	3 February 2002	1263	710	3.6	16	286	1.8
14	Dementia	85,002	8 February 2002	2124	1142	2.1	10	248	1.9
15	Dizziness	6828	1 June 2004	326	221	11.2	2	17	1.5
16	Duchenne muscular dystrophy	45,210	10 September 2004	1627	860	2.2	5	43	1.9

(continued)

Table 7.3 (continued)

Number	Topic title	Page size (Bytes)	Date created	Total number of revisions	Total number of distinct authors	Average time between edits (days)	Average edits per month	Revision in last 12 months	Average edits per user
17	Encephalitis	11,287	25 February 2002	843	541	5.3	12	68	1.6
18	Encephalopathy	7163	7 December 2003	228	161	16.3	0	18	1.4
19	Epilepsy	82,991	20 January 2002	4074	1690	1.1	16	838	2.4
20	Ganglion	4135	28 May 2001	282	202	16.8	5	29	1.4
21	Hearing	15,794	29 September 2003	927	527	4.2	7	85	1.8
22	Lower motor neuron lesion	3089	6 May 2007	51	39	50.4	2	9	1.3
23	Meningitis	72,591	25 February 2002	3390	1573	1.3	4	122	2.2
24	Motor neuron	10,489	29 November 2002	410	247	10.3	6	51	1.7
25	Multiple sclerosis	85,774	20 October 2001	4979	1929	0.9	6	333	2.6
26	Myasthenia gravis	36,120	2 May 2001	981	522	4.9	3	38	1.9
27	Myelin	15,785	28 May 2001	555	353	8.3	20	53	1.6
28	Myopathy	8880	21 February 2005	161	112	20.3	0	4	1.4
29	Neuromuscular junction	29,492	26 May 2004	391	197	9.3	1	24	2.0
30	Neuron	52,220	6 November 2001	2904	1555	1.6	10	177	1.9
31	Olfaction	39,512	16 December 2002	925	568	4.5	4	63	1.6
32	Parkinson's disease	101,655	22 July 2001	6248	2003	0.8	3	73	3.1
33	Peripheral neuropathy	27,879	20 April 2004	509	255	7.3	1	78	2.0

34	Sensory neuron	11,178	27 July 2004	290	175	12.3	1	28	1.7
35	Smell	275	25 November 2002	665	410	6.3	6	22	1.6
36	Spinal cord	30,613	30 January 2002	1382	842	3.3	13	79	1.6
37	Spinal cord injury	62,444	10 October 2004	1062	474	3.3	0	96	2.2
38	Stroke	115,479	16 April 2001	3425	1636	1.4	2	150	2.1
39	Taste	65,554	2 April 2002	2086	1112	2.1	10	114	1.9
40	Upper motor neuron lesion	3615	30 January 2005	81	55	40.7	0	4	1.5
41	Vertigo	28,227	04 September 2005	1258	685	2.5	23	99	1.8
42	Wernicke's area	11,390	21 March 2004	231	176	16	0	23	1.3

these images/illustrations were not at the standards expected of educational resources and the images used were not labelled to explain radiological, microbiological and pathological changes.

As indicated by the Wikipedia administrators, several articles were incomplete. These deficiencies may be summarized as follows: (1) articles in their early stages, for example, the articles on 'peripheral neuropathy', 'brainstem', 'autonomic nervous system', 'Broca's area', 'encephalopathy', 'motor neuron', 'lower motor neuron lesion', 'smell', and 'spinal cord'; (2) articles showing deficiencies in some content or needing tables, images, illustrations or media to make the message meaningful and enhance their educational value, for example, the articles on 'brainstem glioma', 'dizziness', 'encephalopathy', 'myopathy', 'sensory neuron', 'smell', 'upper motor neuron lesion'; and (3) articles requiring the addition of proper citations for some statements, for example, the article on 'sensory neuron', and 'encephalitis'. Although the number of references for the 42 articles was 1517, some articles had no references and a number of problems were identified in the list of references and the quality of references cited. Interestingly, none of the 1517 references was a Wikipedia citation. Recently, Bould et al. (2014) found that 1433 full text articles from 1008 journals indexed in Medline, PubMed or Embase with 2049 Wikipedia citations were accessed. They also found that the frequency of most citations occurred after December 2010. The Wikipedia citations were not limited to journals with a lower or no impact factor, but were in many journals with high impact factor. The authors warned journal editors and peer-reviewers to use caution when publishing articles that cite Wikipedia. The readability of Wikipedia articles was geared at college level indicating that the articles were not written for the public and the language used was suitable for the medical students.

It is obvious from recent research that Wikipedia has continuously worked to improve the quality of its medical/health content (Chiang et al., 2012; Raspberry, 2014). However, few articles meet the quality standards that medical schools would expect before recommending such resources to medical students. These findings have been reached when researchers evaluated Wikipedia 'gastroenterology' and 'hepatology' articles (Azer, 2014). A few researchers reported that Wikipedia articles are useful resources for patients with hand illness (Burgos, Bot, & Ring, 2012), and a reliable source for nephrology patients although written at a college reading level (Thomas, Eng, de Wolff, & Grover, 2013). It was also reported that Wikipedia was a prominent source of online health information compared to other online health information providers (Laurent & Vickers, 2009). Others reported that the quality of osteosarcoma related information in the English version of Wikipedia was inferior to the patient information provided by the US National Cancer Institute (Leithner et al., 2010).

The methods used in evaluating Wikipedia articles in this study aimed at providing a critique of accuracy, clarity, quality, and adequacy of content committed to nervous system articles. Three evaluators conducted the assessment of the Wikipedia articles and the methods were used in earlier publications (e.g., Azer, 2014). The agreement among evaluators had mean \pm SD range of scores of 0.65 \pm 0.10 to 0.79 \pm 0.12.

Wikipedia articles need peer-review by experts and professionals. Harnad (1999) described peer-review as a quality control and certification process to ensure accuracy and validity of material produced in an academic environment. The results from this study show that most articles were updated regularly and the mean \pm SD number of revisions of the 42 articles was 1298 \pm 1418.00 and the average time between edits varied from 0.8 to 70.5 days; 10.15 \pm 13.94 (mean \pm SD). However, anonymous users of Wikipedia made approximately 30 % of edits. Generally, it is difficult to know the actual experience, level of education and skills of the Wikipedia articles.

Several suggestions have been made to improve the quality of editing of Wikipedia articles by doctors (Kint & Hart, 2012), experts and professionals who have specialized in the designated topic. Reavley et al. (2012) suggested that professional associations could create task forces for reviewing Wikipedia and even place an approval statement on acceptable articles. Wicks and Bell (2012) suggested that professional societies could nominate or suggest peer-reviewers that can take such responsibilities.

Recently, 'WikiProject Medicine' has been introduced where people interested in medical and health content on Wikipedia can discuss, collaborate or debate issues. Additionally, Wikipedia articles have also been categorised in regard to their status by administrators. For example, awarding of a golden star means 'Featured Article', or awarding of 'A' means Approved A-Class article etc. Details about Wikipedia categorisation are given on the following link (http://en.wikipedia.org/wiki/Category:FA-Class_medicine_articles). The aim of such categorisation is to help readers and editors/authors understand the relative status and possible veracity of the article.

The study reported in this chapter has several limitations; it evaluated only 42 Wikipedia articles and was limited to English-language topics on the nervous system. Therefore, generalization of these results to other medical or healthcare disciplines is not applicable. More work is needed in the future to evaluate Wikipedia articles on a wider range of medical and surgical diseases.

However, despite these limitations, this study raises important issues in the area of medical education and medical informatics particularly for problem-based learning programmes where self-directed learning is an important domain in the curriculum design (Artino, Cleary, Dong, Hemmer, & Durning, 2014). Expected directions in research in this area may include:

1. Expanding the evaluation of Wikipedia articles to other medical and surgical topics so that a conclusive evaluation of Wikipedia articles could be made.
2. Further assessment of the data provided by Wikipedia in regard to updating and revision of its articles in order to assess the quality of such revisions and understand why, despite recording higher numbers of revisions, articles were not at the standards required for an educational resource.
3. Assessing the impact of engaging medical students in reviewing Wikipedia articles and critically assessing them on their learning and understanding of topics evaluated and studied.

7.5 Conclusion

This is ongoing research; the findings from this study suggest that there were deficiencies and scientific errors in most Wikipedia articles evaluated. Considering the tendency of medical students to depend on Wikipedia in their learning, it may be necessary to educate students in critically engaging with online information by, for example, using guidelines such as the criteria used in this study in evaluating online resources. Given the expectation of medical teachers that students should take responsibilities of their self-regulated learning, Wikipedia articles could be a resource for critical evaluation and content improvement. These recommendations together with the need of medical schools to offer training to its students on how to select their learning resources is necessary.

Acknowledgements This work was funded by the College of Medicine Research Center, Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia. The author thanks Dr. Sarah Azer, and Diana Azer for their help in data processing. He also thank Dr Raju Kumar, Mr. Philip Feeley and Dr. Sharon O’Conner for their help.

References

- Alegría, D., Boscardin, C., Poncelet, A., Mayfield, C., & Wamsley, M. (2014). Using tablets to support self-regulated learning in a longitudinal integrated clerkship. *Medical Education Online*, 19. Retrieved from <http://dx.doi.org/10.3402/meo.v19.23638>
- Andreoli, T. E., Benjamin, I. J., Griggs, R. C., & Ewing, E. J. (2010). *Andreoli and Carpenter’s Cecil essentials of medicine* (8th ed.). Philadelphia, PA: Saunders.
- Artino, A. R., Cleary, T. J., Dong, T., Hemmer, P. A., & Durning, S. J. (2014). Exploring clinical reasoning in novices: A self-regulated learning microanalytic assessment approach. *Medical Education*, 48(3), 280–291. doi:10.1111/medu.12303.
- Azer, S. A. (2014). Evaluation of gastroenterology and hepatology articles on Wikipedia: are they suitable as learning resources for medical students? *European Journal of Gastroenterology & Hepatology*, 26(2), 155–163. doi:10.1097/meg.0000000000000003.
- Bould, M. D., Hladkovicz, E. S., Pigford, A.-A. E., Ufholz, L.-A., Postonogova, T., Shin, E., & Boet, S. (2014). References that anyone can edit: Review of Wikipedia citations in peer reviewed health science literature. *British Medical Journal*, 348, g1585. doi:10.1136/bmj.g1585
- Burgos, C., Bot, A., & Ring, D. (2012). Evaluating the effectiveness of a wiki internet site for medical topics. *Journal of Hand and Microsurgery*, 4(1), 21–24. doi:10.1007/s12593-012-0064-0.
- Chiang, C. D., Lewis, C. L., Wright, M. D. E., Agapova, S., Akers, B., Azad, T. D., Herzog, E. D. (2012). Learning chronobiology by improving Wikipedia. *Journal of Biological Rhythms*, 27(4), 333–336. doi:10.1177/0748730412449578
- Colledge, N. R., Walker, B. R., & Ralston, S. H. (2010). *Davidson’s principles & practice of medicine* (21st ed.). Edinburgh, England: Elsevier, Churchill Livingstone.
- DISCERN Project. (1999). Retrieved from <http://www.discrim.org.uk/index.php>
- English Wikipedia. (2014). Retrieved from http://en.wikipedia.org/wiki/English_Wikipedia
- Grewal, P., & Alagaratnam, S. (2013). The quality and readability of colorectal cancer information on the internet. *International Journal of Surgery*, 11(5), 410–413. doi:10.1016/j.ijso.2013.03.006.

- Grohol, J. M., Slimowicz, J., & Granda, R. (2014). The quality of mental health information commonly searched for on the internet. *Cyberpsychology, Behavior and Social Networking*, 17(4), 216–221. doi:10.1089/cyber.2013.0258.
- Guarino, S., Leopardi, E., Sorrenti, S., De Antoni, E., Catania, A., & Alagaratnam, S. (2014). Internet-based versus traditional teaching and learning methods. *The Clinical Teacher*, 11(6), 449–453. doi:10.1111/tct.12191.
- Harnad, S. (1999). Free at last: The future of peer-reviewed journals. *D-Lib Magazine*, 5(12). Retrieved from <http://www.dlib.org/dlib/december99/12harnad.html>
- Hughes, B., Joshi, I., Lemonde, H., & Wareham, J. (2009). Junior physician's use of Web 2.0 for information seeking and medical education: A qualitative study. *International Journal of Medical Informatics*, 78(10), 645–655. doi:10.1016/j.ijmedinf.2009.04.008.
- Jaffe, J., Tonick, S., & Angell, N. (2014). Quality of Web-based information on epidural anesthesia. *Obstetrics & Gynecology*, 123, 115S. doi:10.1097/01.aog.0000447063.32706.46.
- Kharbanda, A. B., Stevenson, M. D., Macias, C. G., Sinclair, K., Dudley, N. C., Bennett, J., Dayan, P. S. (2012). Interrater reliability of clinical findings in children with possible appendicitis. *Pediatrics*, 129(4), 695–700. doi:10.1542/peds.2011-2037
- Kingsley, K., Galbraith, G. M., Herring, M., Stowers, E., Stewart, T., & Kingsley, K. V. (2011). Why not just Google it? An assessment of information literacy skills in a biomedical science curriculum. *BMC Medical Education*, 11(1), 17. doi:10.1186/1472-6920-11-17.
- Kint, M., & Hart, D. P. (2012). Should clinicians edit Wikipedia to engage a wider world web? *British Medical Journal*, 345, e4275. doi:10.1136/bmj.e4275.
- Kolski, D., Arlt, S., Birk, S., & Heuwieser, W. (2013). Use and acceptance of Wiki systems for students of veterinary medicine. *GMS Zeitschrift für Medizinische Ausbildung*, 30(1), Doc10. doi:10.3205/zma000853.
- Kumar, P., & Clark, M. (2013). *Kumar and Clark's clinical medicine* (8th ed.). Edinburgh, England: Elsevier.
- Lau, A. Y., & Coiera, E. W. (2008). Impact of Web searching and social feedback on consumer decision making: A prospective online experiment. *Journal of Medical Internet Research*, 10(1), e2. doi:10.2196/jmir.963.
- Laurent, M. R., & Vickers, T. J. (2009). Seeking health information online: Does Wikipedia matter? *Journal of the American Medical Informatics Association*, 16(4), 471–479. doi:10.1197/jamia.m3059.
- Leithner, A., Maurer-Ertl, W., Glehr, M., Friesenbichler, J., Leithner, K., & Windhager, R. (2010). Wikipedia and osteosarcoma: A trustworthy patients' information? *Journal of the American Medical Informatics Association*, 17(4), 373–374. doi:10.1136/jamia.2010.004507.
- Longo, D. I., Fauci, A. S., Kasper, D. L., Hauser, S. L., Jameson, J. L., & Loscalzo, J. (2012). *Harrison's principles of internal medicine* (18th ed.). New York, NY: McGraw-Hill.
- Masters, K. (2008). Access to and use of the Internet by South African general practitioners. *International Journal of Medical Informatics*, 77(11), 778–786. doi:10.1016/j.ijmedinf.2008.05.008.
- McPhee, S. J., Papadakis, M. A., & Rabow, M. W. (2011). *Current medical diagnosis & treatment* (50 Anniversary ed.). New York: McGraw-Hill, Lange.
- Patil, A. S., Elnaggar, A. C., Kumar, S., Ling, F. W., Stritter, F. T., Temiyakarn, L., & Tillmanns, T. D. (2014). Guided outcomes in learned efficiency model in clinical medical education: A randomized controlled trial of self-regulated learning. *American Journal of Obstetrics and Gynecology*, 211(5), 544.e1–544.e7. doi:10.1016/j.ajog.2014.05.048
- Petty, J. (2013). Interactive, technology-enhanced self-regulated learning tools in healthcare education: A literature review. *Nurse Education Today*, 33(1), 53–59. doi:10.1016/j.nedt.2012.06.008.
- Prasannan, L., Gabbur, N., & Haughton, M. (2014). Use of Web resources among medical students at a large urban medical center. *Obstetrics & Gynecology*, 123, 118S. doi:10.1097/01.aog.0000447071.47953.34.
- Raspberry, L. (2014). Wikipedia: What it is and why it matters for healthcare. *British Medical Journal*, 348, g2478. doi:10.1136/bmj.g2478.

- Reavley, N. J., Mackinnon, A. J., Morgan, A. J., Alvarez-Jimenez, M., Hetrick, S. E., Killackey, E., Jorm, A. F. (2012). Quality of information sources about mental disorders: A comparison of Wikipedia with centrally controlled web and printed sources. *Psychological Medicine*, 42(8), 1753–1762. doi:[10.1017/s003329171100287x](https://doi.org/10.1017/s003329171100287x)
- Schmidt, H. G., Mamede, S., van den Berge, K., van Gog, T., van Saase, J. L., & Rikers, R. M. (2014). Exposure to media information about a disease can cause doctors to misdiagnose similar-looking clinical cases. *Academic Medicine*, 89(2), 285–291. doi:[10.1097/acm.000000000000107](https://doi.org/10.1097/acm.000000000000107).
- Thomas, G. R., Eng, L., de Wolff, J. F., & Grover, S. C. (2013). An evaluation of Wikipedia as a resource for patient education in nephrology. *Seminars in Dialysis*, 26(2), 159–163. doi:[10.1111/sdi.12059](https://doi.org/10.1111/sdi.12059).
- Tsivgoulis, G., Ribo, M., Rubiera, M., Vasdekis, S. N., Barlinn, K., Athanasiadis, D., Alexandrov, A. V. (2013). Real-time validation of transcranial Doppler criteria in assessing recanalization during intra-arterial procedures for acute ischemic stroke: An international, multicenter study. *Stroke*, 44(2), 394–400. doi:[10.1161/strokeaha.112.675074](https://doi.org/10.1161/strokeaha.112.675074)
- Tyson, T. R. (2000). The Internet: Tomorrows portal to non-traditional health care services. *Journal of Ambulatory Care Management*, 23(2), 1–7. doi:[10.1097/00004479-200004000-00002](https://doi.org/10.1097/00004479-200004000-00002).
- Van der Marel, S., Duijvestein, M., Hardwick, J. C., van den Brink, G. R., Veenendaal, R., Hommes, D. W., & Fidder, H. H. (2009). Quality of web-based information on inflammatory bowel diseases. *Inflammatory Bowel Diseases*, 15(12), 1891–1896. doi:[10.1002/ibd.20976](https://doi.org/10.1002/ibd.20976)
- Vargas, C. R., Chuang, D. J., Ganor, O., & Lee, B. T. (2014). Readability of online patient resources for the operative treatment of breast cancer. *Surgery*, 156(2), 311–318. doi:[10.1016/j.surg.2014.03.004](https://doi.org/10.1016/j.surg.2014.03.004).
- Westbrook, J. I., Gosling, A. S., & Coiera, E. W. (2005). The impact of an online evidence system on confidence in decision making in a controlled setting. *Medical Decision Making*, 25(2), 178–185. doi:[10.1177/0272989x05275155](https://doi.org/10.1177/0272989x05275155).
- Wicks, P., & Bell, V. (2012). Letter to the Editor: Quality of mental health information on Wikipedia. *Psychological Medicine*, 42(04), 891. doi:[10.1017/s0033291712000086](https://doi.org/10.1017/s0033291712000086).