Quality of information sources about mental disorders: a comparison of Wikipedia with centrally-controlled web and printed sources

Nicola J Reavley, Andrew J Mackinnon, Amy J Morgan, Mario Alvarez-Jimenez, Sarah E Hetrick, Eoin Killackey, Barnaby Nelson, Rosemary Purcell, Marie BH Yap, Anthony F Jorm.

Correspondence to:

N J Reavley

Orygen Youth Health Research Centre, Centre for Youth Mental Health

University of Melbourne

Locked Bag 10

Parkville

VIC 3052

Australia

nreavley@unimelb.edu.au

Quality of information sources about mental disorders: a comparison of Wikipedia with centrally-controlled web and printed sources

Abstract

Background

While mental health information on the internet is often of poor quality, relatively little is known about the quality of websites such as Wikipedia, which involve participatory information sharing. The aim of this paper was to explore the quality of user-contributed mental health-related information on Wikipedia and compare this with centrally-controlled information sources.

Method

Content on 10 mental-health-related topics was extracted from 14 frequently-accessed websites (including Wikipedia) providing information about depression and schizophrenia, Encyclopaedia Britannica, and a psychiatry textbook. The content was rated by experts according to the following criteria: accuracy, up-to-dateness, breadth of coverage, referencing and readability.

Results

The results showed that ratings varied significantly between resources according to topic. Across all topics, Wikipedia was the most highly rated in all domains except readability.

Conclusions

The quality of information on depression and schizophrenia on Wikipedia is generally as good as, or better than, that provided by centrally-controlled websites, Encyclopaedia Britannica and a psychiatry textbook.

Introduction

It is estimated that over 1.9 billion people have access to over 312 million sites on the internet (Internet World Statistics, 2011, Netcraft, 2011) and that as many as 80% of internet users in developed countries use the internet to search for information on health problems, symptoms, diseases and treatments (Kummervold *et al.*, 2008, Pew Internet & American Life Project, 2006). Information on mental disorders is commonly accessed online, particularly by those with a psychiatric diagnosis and their supporters or carers (Khazaal *et al.*, 2008, Powell & Clarke, 2006, Ybarra & Suman, 2006).

The growth in health information on the internet has been followed by an increase in the number of studies analysing its quality. In a 2002 review, Eysenbach *et al.* (2002) reported that 55 of 79 such studies considered quality to be a problem, although accuracy varied between health domains, with up to 90% of diet and nutrition information assessed as being unreliable compared to only 5% of that for cancer. A recent review of studies assessing the quality of websites providing information about mental disorders found that most of the research concluded that quality was poor, although site selection and rating methods varied, with some having unknown validity (Reavley & Jorm, 2010).

A relatively recent feature of the debate about the quality of health information on the internet centres on websites that involve users in information sharing and collaboration, rather than viewing them as passive consumers of content created by experts. Often known as 'Web 2.0', this participatory model of web usage is associated with a number of applications, including social networking sites, blogs, media sharing sites and wikis. One of the best known of these is Wikipedia, the online encyclopaedia that anyone can edit. Over 50% of internet users now source information from Wikipedia (Pew Internet & American

Life Project, 2011), which has over 3.3 million English language articles and has become a prominent source of online health information (Laurent & Vickers, 2009). In 2005, a study comparing the quality of science articles in Encyclopaedia Britannica with those in Wikipedia found numerous errors in both, but that the difference in accuracy was not particularly great (Giles, 2005). A number of other studies have explored the quality of health information on Wikipedia, with some concluding that the information was of poor quality, and others reporting that it was of acceptable or even high quality (Heilman *et al.*, 2011).

In this study we explored the quality of the user-contributed mental health-related information on Wikipedia and compared this with information from sources that are centrally controlled, including websites , as well as Encyclopedia Britannica and a comprehensive psychiatry textbook. We systematically examined the quality of information on both a high-prevalence mental disorder (depression) and a low-prevalence severe disorder (schizophrenia).

Methods

Selection of sites and topics

The selection of websites from which material was extracted was based on the top 10 Google search results for either of the terms 'depression' (in March 2010) or 'schizophrenia' (in May 2010). The sites chosen by this method are likely to reflect those encountered by a typical user (Eysenbach & Kohler, 2002). Websites that were portals to the content of other sites were excluded. Six sites appeared in the top 10 search results for both topics, four were unique to depression and four to schizophrenia. Overall 14 websites were selected.

Ten mental-health related topics were chosen, five relating to depression and five to schizophrenia. An attempt was made to choose topics that were relatively specific (to facilitate ease of searching), rapidly evolving (to facilitate assessment of up-to-dateness) and controversial (to facilitate assessment of accuracy and breadth of coverage). The depression topics were: (1) antidepressants and suicide in young people; (2) gambling and depression; (3) side effects of ECT and depression; (4) fish oils for depression; and (5) the relationship between attention deficit hyperactivity disorder (ADHD) and depression. The schizophrenia topics were (1) the relationship between cannabis and psychosis/schizophrenia; (2) childhood onset of psychosis; (3) schizophrenia and violence; (4) side effects of antipsychotics; and (5) stigma and schizophrenia.

Using the topic terms (or synonyms) as key words for the searches or via manual browsing, content relating to these topics was extracted from the selected websites and also from the most recent edition of Kaplan and Sadock's Comprehensive Textbook of Psychiatry (Sadock *et al.*, 2009) and the online version of Encyclopaedia Britannica. Between May and August 2010, content relevant to the search topic (either the whole page or (in the case of very long pages) a section of the page) from each source was extracted by two reviewers working separately. Content was then compared and a consensus reached on the content to be included in the rating assessment. The content for the rating assessments was blinded by removing any information that could identify the source sites. Word counts for the topics are given in Table 1. The order of each source was randomised using the list randomizer at http://www.random.org. Ethical approval was not required.

Participants

An evaluation group was formed comprising three psychologists with clinical and research expertise in depression and three in schizophrenia. These experts rated each of the topics related to depression or schizophrenia respectively.

Source assessment

The content of each website was rated on a five-point scale in the following domains: accuracy, up-to-dateness, breadth of coverage, referencing and readability. The following explicit anchors for points 1, 3 and 5 were used :

- Accuracy: 1 Many errors of fact or unsubstantiated opinions, 3 Some errors of fact or unsubstantiated opinions; 5 All information factually accurate;

- Up-to-dateness: 1 Generally not up-to-date, 3 Information partly up-to-date, 5 All information up-to-date;

- Breadth of coverage: 1 Limited or no coverage of topics, 3 Several topics covered, 5 A broad range of topics covered;

- Referencing: 1 No referencing, 3 Partial referencing of statements or referencing with secondary sources, 5 Statements are consistently referenced;

- Readability: 1 Readability suitable for someone with university education, 3 Readability suitable for someone who has completed secondary education, 5 Readability suitable for someone with some secondary education.

Initially, raters met as a group to discuss rating procedures. A 'pilot' rating exercise was then undertaken with a subsequent meeting of the group to discuss and resolve differences of opinion. After the rating was completed and agreement between the raters assessed (see below), domains for which the mean intraclass correlation coefficients (ICCs) fell below 0.5

were noted. For these domains, raters were asked to meet and come to a consensus on the ratings. Agreement was re-evaluated after this process.

Readability was also assessed using the Flesch-Kincaid Grade Level Index, an objective measure of the level of reading difficulty of text, which is scaled to reflect the number of years of education required to read the text. The index reflects sentence length and word complexity (number of syllables) (Kincaid *et al.*, 1975). The index was calculated for each topic from each source using the Readability Calculator at http://www.online-utility.org.

Statistical analysis

Agreement between the three raters was assessed for ratings of each topic in each domain. ICCs were calculated for the average of the ratings using a mixed effects model (McGraw & Wong, 1996). Differences between resources within each domain were investigated using mixed models ANOVA with resource and topic as fixed factors and rater as a random factor. To assist interpretation, a pseudo R^2 value was calculated for each factor in the model. This index was calculated as the residual variance reduction that resulted from adding each term to the model as a proportion of the residual variance of the model including only an intercept term (Singer & Willett, 2003). Unlike the *F* tests reported, this index may be sensitive to the order in which terms are added to the model. However, reversing the introduction of 'Information Source' and 'Topic' made almost no difference to the results.

Because our interest was in the quality of an information source as a whole, rather than individual topics in a source, interpretation focused on the main effect of source. Sources were ordered by average ratings across all domains and topics.

Results

Interrater reliability

Mean, minimum and maximum intraclass correlations for average ratings for the schizophrenia topics in each domain were as follows: Accuracy: 0.82 (0.76-0.89); Breadth: 0.59 (0.25-0.87); Up-to-dateness: 0.83 (0.73-0.92); Referencing: 0.84 (0.72-0.95); and Readability: 0.69 (0.60-0.78). For the schizophrenia topics, with the exception of ratings of breadth, agreement was high and statistically significant for all ratings. Agreement regarding breadth of coverage for side effects of antipsychotics and for cannabis and psychosis/schizophrenia were notably lower than for other topics (ICC=0.43, p=0.106 and ICC=0.25, p=0.205, respectively).

Mean, minimum and maximum intraclass correlations for average ratings for the depression topics in each domain were as follows: Accuracy: 0.59 (0.20-0.82); Breadth: 0.88 (0.74-0.96); Up-to-dateness: 0.75 (0.60-0.91); Referencing: 0.89 (0.83-0.94); and Readability: 0.87 (0.83-0.90). For the depression topics, with the exception of ratings of accuracy, agreement was high and statistically significant for all ratings. The low ICC regarding the accuracy of material on the topic of gambling and depression (ICC=0.20, p=0.254) was due to low between resource variation rather than poor absolute agreement, as reviewers agreed completely for most resources and differed by no more than a point for others.

Expert quality ratings

In all domains, quality varied significantly between sources according to topic, but the strongest effects were between sources (see Tables 2 and 3). Generally, greater differences

between ratings of resources were observed for schizophrenia than for depression, with notable diversity of ratings for individual topics in particular sources.

Schizophrenia ratings

Figure 1 depicts average ratings for each resource in each domain for schizophrenia (along with the minimum and maximum average ratings for each topic in each domain). Wikipedia received the highest ratings for accuracy and was rated consistently for all topics. Accuracy was rated as being at least 'average' for all resources on most topics. Most resources were rated around the average level on breadth of coverage, with the Kaplan and Sadock textbook receiving the highest ratings. Some resources showed substantial variability in ratings of breadth across different topics.

For up-to-dateness, most sources were rated in the average to good range, with two (Mentalhealth.com and Encyclopaedia Britannica) being rated as poorer. Kaplan & Sadock was rated amongst the best sources in this regard, and consistently so across topics. Very few sources were well rated on referencing, although ratings for some topics were 'average'. Wikipedia was clearly the most highly rated on this domain. Readability for many sources was rated as above average. Wikipedia and Kaplan & Sadock received the poorest ratings.

Depression ratings

Figure 2 depicts average ratings for each source in each domain for depression (along with the minimum and maximum average ratings for each topic in each domain). For accuracy, compared to other domains rated, there was comparatively little variation between sources and topics. As with the schizophrenia topics, Wikipedia was rated highest on average, although this website had comparatively large variation across different topics. Wikipedia,

National Institute of Mental Health (nimh.nih.gov), webmd.com, and Kaplan & Sadock were rated as having above average breadth of coverage within the topics studied, while depression.com and National Health Service (nhs.uk) had poor coverage. Other resources fell in the intermediate band.

Wikipedia was clearly the most highly rated resource on the domain of up-to-dateness. Resources varied substantially in their level of referencing, with many providing few or no citations of the medical literature. Ratings were relatively consistent across topics. It is notable that a number of online resources, notably Wikipedia and NIMH, were rated as comparable to, or better than, Kaplan & Sadock. Rated readability varied widely between resources and to some extent, negatively mirrored referencing, in that resources with fewer references were rated as being more readable and vice versa. There were exceptions to this pattern, with Kaplan & Sadock being rated as the least readable resource. Of the online resources, Wikipedia was rated the least readable, although some of its topics received an average rating.

Flesch-Kincaid Grade Level Indices

Figures 3 and 4 show the Flesch-Kincaid Grade Level averaged over topics for each information source. For depression sources, the textbook was evaluated as requiring tertiary levels of education to read. This is perhaps not surprising given the intended audience of the book. However, five other sources were evaluated as requiring higher levels of education than completion of secondary schooling to be read effectively. Amongst these was Wikipedia. The reading level for Encyclopedia Britannica was comparably high. Only three sources had average levels clearly less than high school completion.

Results for schizophrenia sources were similar to depression with the textbook, Wikipedia and Encyclopaedia Britannica having high scores. Average reading levels were slightly higher than for depression, with only two resources—WebMD and Mentalhealth.com—having an average level of 12 years of education (equivalent to high school completion) or below.

Discussion

The quality of information about depression and schizophrenia on Wikipedia was generally rated higher than other centrally-controlled resources, including 14 mental health-related websites, Encyclopaedia Britannica and Kaplan & Sadock's Comprehensive Textbook of Psychiatry. These findings may help to answer one of the most commonly raised concerns about collaboratively-created websites, namely how 'good' is the information found there? In the case of information about topics relating to depression and schizophrenia, particularly those that are relatively controversial and rapidly evolving, the answer appears to be that the quality is relatively high, as rated by experts in the field. These findings largely parallel those of other recent studies of the quality of health information on Wikipedia, including those that have assessed the quality of information on drugs (Clauson *et al.*, 2008), surgical procedures (Devgan *et al.*, 2007), for medical students (Pender *et al.*, 2009), nursing students (Haigh, 2010), for use in a laboratory observations database (Friedlin & McDonald, 2010), on gastroenterological conditions (Czarnecka-Kujawa *et al.*, 2008), cancer (Rajagopalan *et al.*, 2010) and pathology informatics (Kim *et al.*, 2010). Despite variability in these studies' methodologies and conclusions, the overall implication is that Wikipedia articles on health topics typically contain relatively few factual errors, although they may

lack breadth of coverage. They are also generally well-referenced, though not always easy to understand (Heilman *et al.*, 2011).

In rapidly-evolving fields such as health, a potential strength of web-based information is the ease of updating information offered by this platform. This has led to the claim that traditional peer-reviewed medical articles may be made obsolete by the advent of Wikipedia (Frishauf, 2006). As might be expected, Wikipedia was the most highly rated source on the domain of up-to-dateness. However, it is interesting to note that most online sources did not eclipse the rating achieved by the Kaplan & Sadock textbook (which is typically updated every 4-5 years), although there was considerable variability across topics. This suggests that many centrally-controlled websites do not exploit opportunities to update information, or they may not have the required resources to do so. Consistent with this conclusion, a recent trial found that assessment of the quality of website information and feedback to web administrators did not lead to improvement (Jorm *et al.*, 2010).

There are several limitations to this study, including the extent to which some of the ratings are subjective and may be subject to bias, particularly as the raters were working at the same institution. However, this limitation may be considered in the broader context of the issue of expert rating of the quality of scientific information, including that of peer review, which, while widely used, is generally considered to have limited evidence of validity (Jefferson *et al.*, 2002). In addition, the large variability of coverage between topics, which was a feature of the better-rated resources, may limit conclusions regarding overall site quality. In addition, care must be exercised in interpreting the absolute values of the Flesch-Kincaid Grade Level indices as it was developed and has been evaluated in a different context to medical communication. Further, the topics covered require use of long,

multisyllabic words to which the index is sensitive. However, it is clear that most of the resources make reading demands that would exceed the capacity of many users. None had reading levels consistent with primary completion/early secondary school level, despite approximately half of those in many developed countries having a reading age equivalent to primary school completion (National Work Group on Literacy and Health, 1998, Office for National Statistics, 1996) Few, if any, would meet criteria for formal patient information material or plain language statements for trial participant recruitment (Paasche-Orlow *et al.*, 2003). Further research should aim to discover how such information affects consumer health behaviours such as help seeking and use of evidence-based treatments. Such research might involve naturalistic reports of user behaviour (Frost *et al.*, 2008, Sillence *et al.*, 2007) and may be assisted by the web's move towards greater interactivity, information sharing, and collaboration. A further limitation involves the comparison of the 2009 version of the Kaplan & Sadock textbook (which is unlikely to contain references to anything published after 2008 at the latest) with websites examined in 2010, which could contain later references. However, there is some evidence that while websites containing health information have the potential to be constantly updated with new information, they are in fact relatively unlikely to change over time periods of one or two years (Coquard *et al.*, 2011, Jorm *et al.*, 2010).

Despite these limitations, it appears that the participatory model of web usage and information dissemination, as exemplified by Wikipedia, does generate high-quality information about mental disorders such as depression and schizophrenia. Given the number of patients, would-be patients and concerned others using the internet to search for information on health issues, it appears that Wikipedia is an appropriate

recommendation as an information source. The value of participatory sites could be further enhanced by active contributions by psychologists and members of the medical professions. Some professional organisations, such as the Association for Psychological Science, are now urging their members to contribute to wikis to improve content (Banaji, 2011) and it may even be argued that these professional associations should create task forces to add official statements to Wikipedia entries relevant to the field.

References

Banaji, M. (2011). Harnessing the Power of Wikipedia for Scientific Psychology: A Call to Action. In *Observer*. Association for Psychological Science: Washington DC.

Clauson, K. A., Polen, H. H., Boulos, M. N. & Dzenowagis, J. H. (2008). Scope, completeness, and accuracy of drug information in Wikipedia. *Ann Pharmacother* **42**, 1814-21.

Coquard, O., Fernandez, S., Zullino, D. & Khazaal, Y. (2011). A follow-up study on the quality of alcohol dependence-related information on the web. *Subst Abuse Treat Prev Policy* **6**, 13.

Czarnecka-Kujawa, K., Abdalian, R. & Grover, S. C. (2008). The quality of Open Access and Open Source Internet material in gastroenterology: is Wikipedia appropriate for knowledge transfer to patients? *Gastroenterology* **134**, A-325-A-326.

Devgan, L., Powe, N., Blakey, B. & Makary, M. (2007). Wiki-Surgery? Internal validity of Wikipedia as a medical and surgical reference. *J Am Coll Surg* **205**, S76-S77.

Eysenbach, G. & Kohler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ* **324**, 573-7.

Eysenbach, G., Powell, J., Kuss, O. & Sa, E. R. (2002). Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* **287**, 2691-700.

Friedlin, J. & McDonald, C. J. (2010). An evaluation of medical knowledge contained in Wikipedia and its use in the LOINC database. *J Am Med Inform Assoc* **17**, 283-7.

Frishauf, P. (2006). Are traditional peer-reviewed medical articles obsolete? *MedGenMed* **8**, 5.

Frost, J. H., Massagli, M. P., Wicks, P. & Heywood, J. (2008). How the Social Web Supports patient experimentation with a new therapy: The demand for patient-controlled and patient-centered informatics. *AMIA Annu Symp Proc*, 217-21.

Giles, J. (2005). Internet encyclopaedias go head to head. *Nature* **438**, 900-901.

Haigh, C. A. (2010). Wikipedia as an evidence source for nursing and healthcare students. *Nurse Educ Today* **31**, 135-9.

Heilman, J. M., Kemmann, E., Bonert, M., Chatterjee, A., Ragar, B., Beards, G. M., Iberri, D. J., Harvey, M., Thomas, B., Stomp, W., Martone, M. F., Lodge, D. J., Vondracek, A., de Wolff, J. F., Liber, C., Grover, S. C., Vickers, T. J., Mesko, B. & Laurent, M. R. (2011). Wikipedia: a key tool for global public health promotion. *J Med Internet Res* **13**, e14.

Internet World Statistics (2011). Internet Usage Statistics. Miniwatts Marketing Group.

Jefferson, T., Wager, E. & Davidoff, F. (2002). Measuring the quality of editorial peer review. *JAMA* **287**, 2786-90.

Jorm, A. F., Fischer, J. A. & Oh, E. (2010). Effect of feedback on the quality of suicide prevention websites: randomised controlled trial. *Br J Psychiatry* **197**, 73-4.

Khazaal, Y., Chatton, A., Cochand, S., Hoch, A., Khankarli, M. B., Khan, R. & Zullino, D. F. (2008). Internet use by patients with psychiatric disorders in search for general and medical informations. *Psychiatr Q* **79**, 301-9.

Kim, J. Y., Gudewicz, T. M., Dighe, A. S. & Gilbertson, J. R. (2010). The pathology informatics curriculum wiki: Harnessing the power of user-generated content. *J Pathol Inform* **1**.

Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L. & Chissom, B. S. (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. Naval Technical Training, U. S. Naval Air Station: Memphis, TN.

Kummervold, P. E., Chronaki, C. E., Lausen, B., Prokosch, H. U., Rasmussen, J., Santana, S., Staniszewski, A. & Wangberg, S. C. (2008). eHealth trends in Europe 2005-2007: a population-based survey. *J Med Internet Res* **10**, e42.

Laurent, M. R. & Vickers, T. J. (2009). Seeking health information online: does Wikipedia matter? *J Am Med Inform Assoc* **16**, 471-9.

McGraw, K. O. & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods* **1**, 30-46.

National Work Group on Literacy and Health (1998). Communicating with patients who have limited literacy skills. Report of the National Work Group on Literacy and Health. *J Fam Pract* **46**, 168-76.

Netcraft (2011). April 2011 Web Server Survey.

Office for National Statistics (1996). Adult Literacy Survey: Literacy Level of Adults by Gender and Age. Office for National Statistics: London, UK.

Paasche-Orlow, M. K., Taylor, H. A. & Brancati, F. L. (2003). Readability standards for informed-consent forms as compared with actual readability. *N Engl J Med* **348**, 721-6.

Pender, M. P., Lasserre, K. E., Del Mar, C., Kruesi, L. & Anuradha, S. (2009). Is Wikipedia unsuitable as a clinical information resource for medical students? *Med Teach* **31**, 1095-6.

Pew Internet & American Life Project (2006). Online Health Search. Pew Internet: Washington DC.

Pew Internet & American Life Project (2011). Wikipedia: past and present. Pew Internet: Washington DC.

Powell, J. & Clarke, A. (2006). Internet information-seeking in mental health: population survey. *Br J Psychiatry* **189**, 273-7.

Rajagopalan, M. S., Khanna, V., Stott, M., Leiter, Y., Showalter, T. N. & Dicker, A. (2010). Accuracy of cancer information on the Internet: a comparison of a Wiki with a professionally maintained database. *J Clin Oncol* **7(suppl)**.

Reavley, N. J. & Jorm, A. F. (2010). The quality of mental disorder information websites: A review. *Patient Educ Couns* **in press**.

Sadock, B. J., Sadock, V. A., Ruiz, P. & Kaplan, H. I. (2009). *Kaplan and Sadock's comprehensive textbook of psychiatry*. Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia.

Sillence, E., Briggs, P., Harris, P. R. & Fishwick, L. (2007). How do patients evaluate and make use of online health information? *Soc Sci Med* **64**, 1853-62.

Singer, J. D. & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press: Oxford.

Ybarra, M. L. & Suman, M. (2006). Help seeking behavior and the Internet: a national survey. *Int J Med Inform* **75**, 29-41.

Tables

Table 1 Word counts for topics

Resource	Word count				
Schizophrenia topics	Relationship between cannabis and psychosis/ schizophrenia	Childhood onset of psychosis	Schizophrenia and violence	Side effects of anti-psychotics	Stigma and schizophrenia
Wikipedia	2571	695	348	6938	172
Mayo Clinic	631	3598	604	4148	1763
Kaplan & Sadock	1703	6766	23608	39159	10905
NIMH	225	4666	1374	7628	3020
WebMD	5204	85	2975	15885	3864
schizophrenia.co	11877	7689	10678	15893	1830

m					
eMedicineHealth	909	425	195	2755	0
NHS	5327	402	366	5082	1825
NAMI	2722	1181	5522	4935	9996
MedicineNet	2002	2869	700	6159	3407
Mentalhealth.com	1036	226	1170	0	12979
Encyclopedia Britannica	959	179	921	3789	0

Table 1(cont) Word counts for topics

Resource   Word count

Depression topics	Anti-depressants and suicide in young people	Gambling and depression	Side effects of ECT and depression	Fish oils for depression	Relationship between ADHD and depression
Wikipedia	1857	216	7306	2110	1347
NIMH	2232	0	2914	2492	15345
WebMD	13387	5936	11656	7688	18809
Kaplan & Sadock	4575	1503	28669	3903	10453
Mayo Clinic	1914	809	2815	2131	2792
MedicineNet	7878	5841	3378	5015	3873
helpguide.org	619	189	0	3253	3786
kidshealth.org	1695	0	0	0	3360
beyondblue	2465	2004	401	528	243
Encyclopedia Britannica	0	0	1586	0	0
depression.com	384	0	191	0	0
NHS	1035	361	976	167	1189

Table 2 Mixed model ANOVA of ratings of schizophrenia information for five domains by resource and topic

Domain	Pseudo R^2, (F test[†] , p value)		
	Resource	Topic	Resource×Topic
Accuracy	0.33	0.03	0.28
	(F=16.4, P=0.000)	(F=4.9, P=0.001)	(F=4.0, P=0.000)
Breadth	0.15	0.11	0.27
	(F=6.8, P=0.000)	(F=10.9, P=0.000)	(F=3.2, P=0.000)
Up-to-dateness	0.29	0.02	0.33
	(F=14.9, P=0.000)	(F=4.0, P=0.004)	(F=4.5, P=0.000)
Readability	0.46	0.05	0.22
	(F=13.3, P=0.000)	(F=2.8, P=0.029)	(F=1.6, P=0.024)
Referencing	0.40	0.02	0.08
	(F=27.6, P=0.000)	(F=8.6, P=0.000)	(F=3.9, P=0.000)

† F tests have 11,112; 4,112 and 41,112 degrees of freedom respectively.

Table 3 Mixed model ANOVA of ratings of depression information for five domains by resource and topic

Domain	Pseudo R^2, (F test[†], p value)		
	Resource	**Topic**	**Resource×Topic**
Accuracy	0.23 (F=7.6, P=0.000)	0.04 (F=4.1, P=0.004)	0.17 (F=2.2, P=0.002)
Breadth of coverage	0.33 (F=23.9, P=0.000)	0.14 (F=21.4, P=0.000)	0.30 (F=5.9, P=0.000)
Up-to-dateness	0.26 (F=7.9, P=0.000)	0.00 (F=1.3, P=0.274)	0.22 (F=2.6, P=0.000)
Readability	0.61 (F=23.7, P=0.000)	0.01 (F=1.4, P=0.245)	0.14 (F=2.1, P=0.003)
Referencing	0.58 (F=34.6, P=0.000)	0.00 (F=3.1, P=0.019)	0.09 (F=3.3, P=0.000)

† F tests have 11,93–94; 4,93–94 and 33,93–94 degrees of freedom respectively.