

Review of the quality of information on bipolar disorder on the Internet

Caryl Barnes, Robin Harvey, Alex Wilde, Dusan Hadzi-Pavlovic, Kay Wilhelm, Philip B. Mitchell

Objectives: The aim of the present study was to assess the quality of commonly found websites on bipolar disorder. A specifically designed quality tool, the Bipolar Website Quality Checklist (BWQC), was developed for this purpose.

Methods: The BWQC was developed from quality criteria identified by a literature review of Medline (1966+), Medline in-process and non-indexed citations, PsychINFO, CINAHL, EMBASE, Pub Med, Science citation index, and PsychArticles, using keywords: 'quality, reliability, accuracy, readability, evaluation, assessment, information, internet, web, www'. To identify relevant websites, seven common search engines were accessed and searched using a string of key words: 'bipolar disorder+mania+ manic depression+hypomania'. The top active 15 sites identified were rated by three independent raters, using the BWQC and DISCERN instruments.

Results: There was a wide variability in the quality of the websites reviewed. The Black Dog Institute website was ranked first by the BWQC and DISCERN instruments. The National Institute of Mental Health website was ranked second by DISCERN and seventh by BWQC. The BWQC demonstrated high interrater reliability ($r=0.89$) and correlated strongly ($r=0.78$, $p=0.001$) with the more generic DISCERN instrument. Websites with an editorial board or affiliation to a professional organization or which contained information on a variety of mental health issues had higher quality information on bipolar disorder and its treatment than websites that did not share these characteristics.

Conclusions: High-quality information on bipolar disorder does exist on the Internet. It is important that clinicians are familiar with such websites so that they can recommend the most appropriate site that meets the specific need of the individual. Use of such websites can assist clinicians in adhering to clinical practice guidelines by providing material to augment psychoeducational interventions.

Key words: bipolar disorder, evaluation, quality, websites.

Australian and New Zealand Journal of Psychiatry 2009; 43:934–945

Caryl Barnes, Consultant Psychiatrist, MD Student (Correspondence)
School of Psychiatry, University of New South Wales, Black Dog Institute Building, Hospital Road, Prince of Wales Hospital, Randwick, NSW 2031, Australia. Email: caryl.barnes@awp.nhs.uk

Robin Harvey, Clinical Psychologist

School of Psychology, University of Western Australia, Perth, Western Australia, Australia

Alex Wilde, PhD Student; Dusan Hadzi-Pavlovic, Senior Hospital Scientist; Kay Wilhelm, Conjoint Professor

School of Psychiatry, University of New South Wales, Sydney, New South Wales, Australia; Black Dog Institute, Sydney, New South Wales, Australia

Philip B. Mitchell, Scientia Professor

School of Psychiatry, University of New South Wales, Sydney, New South Wales, Australia; Black Dog Institute, Sydney, New South Wales, Australia; Brain Sciences UNSW, Sydney, New South Wales, Australia

Received 3 March 2009; accepted 28 April 2009.

The convenience and privacy of the Internet makes this an attractive channel for providing and accessing health information [1]. As of March 2009, there were 251.3 million Internet users in North America, a penetration of 74.4% of the total population. Data for Oceania/Australia for the same time found that there were 20.8 million Internet users, 60.5% penetration of the total population [2]. US studies have shown that 80% of these search the Internet each year for health-related information, the equivalent of 95 million adults [3], with 70% reporting that this information influences their decision about treatment [4]. Although there are no published data on how frequently those with bipolar disorder access the Internet, it is likely to be a common tool for gathering information on this condition. Clinically, the provision of psychoeducation on bipolar disorder is now recognized as an important evidence-based component of management [5,6]. The significant role of psychoeducation is also reflected in its prominence in several clinical practice guidelines [7,8].

A major concern in the management of bipolar disorder is the delay in obtaining treatment after the onset of mood disturbance [9]. Such a delay in treatment may reduce the likelihood of response. Because the Internet is now such an accessible means of providing health material, such as information on the clinical characteristics and treatment of bipolar disorder, it has the potential to empower individuals to seek earlier medical review of diagnosis and treatment.

Despite its enormous potential, the quality of general health information on the Internet has been of concern for some time. A systematic review by Eysenbach *et al.* evaluated 79 studies that assessed the quality of health information on the web [10]. Seventy per cent of these studies judged the quality of information to be poor, 22% judged it neutral, and only 9% deemed it to be positive. Although the need to evaluate the quality of health information on the Internet is not contested, there is uncertainty as to how evaluation should be carried out. In 1997 Silberg *et al.* proposed that a set of transparency criteria developed for the print media could be useful indicators of the quality of web-based health information [11]. These were: authorship, attribution, disclosure and currency. In the review by Eysenbach *et al.* the 79 studies had evaluated 5941 different health websites, using 86 different quality criteria [10]. Eysenbach *et al.* distilled these 86 into five main quality measures: (i) accuracy, defined as the 'degree

of concordance of the information provided with best evidence or with generally accepted medical practice'; (ii) completeness, defined as 'the scope or coverage of information contained in the website'; (iii) readability (this was not defined as such by Eysenbach *et al.* but this criterion can be assessed by using formulae such as the Flesch–Kincaid grade level index to give an overall reading level for content; such formulas do not consider important subjective elements such as the use of jargon or tone/mood and writing style that impact on the understandability of material); (iv) design (aesthetics; although criticized by the authors as being too subjective a measure to be useful, design criteria focus on the visual aspect or layout of the site); and (v) technical criteria, defined by Eysenbach *et al.* as 'general, domain-independent criteria, that is, how information is presented'. Eysenbach *et al.* proposed that technical criteria were variations on the Silberg *et al.* transparency criteria. Examples include disclosure of authorship, date of website creation or page update, references provided, author's credentials, email address and details of sponsorship.

In 2000 Griffiths and Christensen published one of the few studies that has specifically examined the quality of mental health information on the Internet [12]. Their cross-sectional survey of websites that provide information on the treatment of depression found the quality of such information to be poor. They questioned the usefulness of the Silberg *et al.* accountability criteria, because they found other website characteristics, such as ownership by an organization and the existence of an editorial board, to be better indicators of content quality. To date there has been only one brief report published that has examined the quality of websites on bipolar disorder [13].

Aims

A survey of websites on bipolar affective disorder was undertaken with the aim of providing guidance on the quality of sites for both patients and mental health clinicians. The specific aims of the present study were to (i) identify major bipolar disorder websites using common search engines; (ii) review the quality of these sites using defined quality criteria; and (iii) develop a new website quality assessment methodology for bipolar disorder sites assessing features not included in current more generic scales.

Methods

Measures of website quality

A literature review was undertaken to identify instruments that would be suitable for evaluating health information websites such as those focused on bipolar disorder. The databases searched included Medline (1966+), Medline in-process and non-indexed citations, PsychINFO, CINAHL, EMBASE, PubMed, Science citation index, and PsychArticles, as well as the Internet search engine Google Scholar. The following key words were used: 'quality' OR 'reliability' OR 'accuracy' OR 'readability' OR 'evaluation' OR 'assessment' AND 'information' AND 'internet' OR 'web' OR 'www'.

Despite the large body of conceptual work that exists in the area of website evaluation, no validated instrument was found that met all the requirements of this review. The most relevant was the DISCERN instrument, a validated tool developed by an expert panel in conjunction with patients (funded by the British Library) to evaluate online information about medical treatment [14]. It was designed to be a generic tool to assist in the evaluation of all health websites, and was not developed to evaluate the quality of specific disorder or treatment content. DISCERN consists of 16 items, the last of which provides an overall global quality score (which was not used in the present review). Items 1–15 are summed to give the total DISCERN score. Each item consists of a 5-point Likert scale with defined anchor points. Given the lack of other validated instruments this tool was used to assist in evaluation of the selected websites. This instrument is now available online (<http://www.discern.org.uk/>). At the time of the present review this online version was not available and so the original paper-based version was used.

In addition, the authors developed a new Bipolar Website Quality Checklist (BWQC) to specifically evaluate the quality of bipolar disorder websites, using the criteria deemed to be the most critical in the bipolar disorder literature. The psychometric properties are described in the present article. All identified websites were evaluated using both the DISCERN and BWQC instruments.

The BWQC (Appendix 1) consists of six subscales, with a total of 56 items: Credibility (seven items); Currency (two items); Objectivity (six items); Availability and Usability (four items); Design and Aesthetics (two items); and Breadth and Accuracy (35 items). These subscales consist mainly of the quality criteria identified by Silberg *et al.* [11] and Eysenbach *et al.* [10], with the exception of the excessively complex Readability criterion. This was excluded because it required the use of specific tools that were unfamiliar to the raters and required specific training to ensure validity. The last subscale (Breadth and Accuracy) consists of 35 items, each of which relates to a specific diagnostic or treatment recommendation in the clinical practice guidelines for the treatment of bipolar disorder published by the Royal Australian and New Zealand College of Psychiatrists [7]. The majority of these recommendations are consistent with those of the clinical practice guidelines of the American Psychiatric Association [8]. This approach was based on that used by Griffiths and Christensen, who used the guidelines for the treatment of depression written by the US Agency for Health

Care Policy and Research in development of their guideline score [15,16]. Their guideline score was determined by the number of times a guideline was accurately reflected by the content of the website. The individual items of the BWQC are detailed in Appendix 1. Each item of the BWQC was developed as a 5-point Likert scale with the following anchors: 0=no, 3=partially and 5=yes. To enhance interrater reliability, the main author (CB) developed a BWQC User Guide that provided detailed descriptors for all items.

Selection of websites

The aim of the present study was to identify websites commonly encountered by typical web users searching for information on bipolar disorder. Seven common search engines were used: www.google.com (both worldwide and Australian versions), www.yahoo, www.altavista, www.ask.com, www.excite.com and www.lycos.com. The search was performed during a 2 week period in August 2007. The key words or terms 'bipolar disorder', 'mania', 'manic depression', and 'hypomania' were entered as a string into the search box for each search engine. The top 20 sites listed by each search engine were recorded and compared. This cut-off was chosen because it has been reported that most people do not investigate beyond this point [10]. Websites that were inactive or did not contain a sufficient amount of information on bipolar disorder (defined a priori as four or more pages) were excluded. Each website was ranked based on both the frequency that it appeared across the seven search engines and the total number of times it was listed. Only websites appearing in at least two search engines were retained, leaving a listing of 16 sites (Table 1). One website became inactive during the study and was therefore excluded, leaving a final total of 15 sites. These were assessed in detail by the three raters over a 3 month period.

Website characteristics

One of the aims of this survey was for the results to be comparable with other studies evaluating mental health websites. Some of the website characteristics used by Griffiths and Christensen [12] in their study of depression websites were therefore adopted: (i) presence of an editorial board; (ii) ownership type (professional, patient or commercial); and (iii) the scope of information. We further subcategorized the latter as 'bipolar disorder information only' and 'bipolar plus', that is, when the website also contained other non-bipolar disorder health information.

Raters

Three raters were used to evaluate the websites. Two (CB and AW) were doctoral students in the School of Psychiatry at the University of New South Wales, which is affiliated with the Black Dog Institute in Sydney; one was a psychiatrist and the other a medical journalist with a science degree. The third rater was a specialist child and adolescent developmental clinical psychologist

in Perth who had previously worked with a commercial health service provider with specialisation in online disease management systems (Sentiens). All raters evaluated these sites independently of each other.

Statistical analyses

The statistical software SPSS version 15 (SPSS, Chicago, IL, USA) was used for the analyses. The mean scores of the three raters for the total DISCERN score and the total and subscale scores of the BWQC were compared against the three chosen website characteristics of Griffiths and Christensen (presence of an editorial board, ownership type, and scope of information). Data were assessed for normality using the Kolmogorov–Smirnov statistic. Those subscales with a normal distribution were compared against the site characteristics using independent t-tests and ANOVA. For subscales with a non-normal distribution, non-parametric statistical tests were used (Mann–Whitney and Kruskal–Wallis). To correct for Type I errors due to multiple testing, the Bonferroni correction method was used. This involved calculating a new alpha by dividing 0.05 by the number of dependent variables that were being evaluated against the site characteristics. ($0.05/8$, $p=0.006$). Scheffé post-hoc tests were used with the ANOVAs. Associations were examined using the Pearson product-moment correlation coefficient. Interrater reliability between the three raters with the BWQC was examined using intra-class correlation coefficients and calculated using the two-way mixed model with absolute agreement type. The authors of the DISCERN instrument used weighted κ as a measure of agreement between each rater. It was felt by these authors to be the ‘appropriate measure for the analysis of data in ordered categories, such as the 5-point Likert scale used to rate each item on the DISCERN as it does not treat all disagreements equally’ [14].

Results

Website selection

Table 1 lists the top 15 websites identified by this process, ranked by both the number of search engines in which they were listed and the total number of hits for that website. The overall total number of hits for each website using the key words previously listed is shown in Table 2.

Website characteristics

The majority of the websites (73%) originated from North America. Sixty per cent of the websites were commercial, with professional and patient websites making up the remaining 40%. The majority (80%) of the websites were owned by an identifiable organization, with one directly owned by a pharmaceutical company. Just under a half of the websites indicated that they had an editorial board (47%). The majority of the websites (67%) could be defined as being bipolar plus, that is, that they contained information about more health issues than only bipolar disorder.

Interrater reliability

The interrater reliability between the three raters for the total DISCERN score was $r=0.61$ ($p=0.009$). The intra-class correlation coefficient for the mean total BWQC score was higher at $r=0.89$ ($p<0.00005$). For the BWQC subscales, there was a high interrater reliability for all scales, with the highest ($r=0.96$; $p<0.0005$) being for Mean Currency, and the lowest ($r=0.75$; $p=0.001$) being for Mean Availability and Usability and Mean Design and Aesthetics. The higher interrater reliability for the BWQC compared to the DISCERN may have related to the intensity of

Table 1. Commonly found websites on bipolar disorder as identified during August 2007

Website address	Rank [†]	No. search engines (n = 7)	Total no. hits
www.helpguide.org	1	5	14
www.pendulum.org	2	5	7
www.nimh.nih.gov	2	5	7
www.en.wikipedia.org	4	5	6
www.bipolar.about.com	5	4	12
www.bipolarworld.net	6	4	4
www.answers.com	6	4	4
www.bipolar.com	6	4	4
www.nlm.nih.gov	9	3	3
www.manic-depression.net	9	3	3
www.mentalhealthchannel.net	9	3	3
www.medicinenet.com	9	3	3
www.netdoctor.co.uk	14	2	3
www.blackdoginstitute.org.au	14	2	3
www.patient.co.uk	14	2	3

[†]Rank based on the frequency the website appeared across the seven search engines as well as the total number of times the website was listed using search keywords.

Table 2. Total number of hits per search engine for keywords 'bipolar disorder + mania + manic depression + hypomania'

Search engine	Total no. hits
Ask.com	174
Excite.com	73
Google (Australia only)	54
Google (world-wide)	9,670
Altavista	15,100
Yahoo	14,900
Looksmart	300

training in the two scales. Both the BWQC and the DISCERN instruments had user guidelines, but rater training was more intense with the new scale. The interrater reliability of the DISCERN was comparable to those scores reported in its initial development. Charnock *et al.* reported that the chance corrected agreement (weighted κ) for the overall quality rating was $\kappa=0.53$ (95% confidence interval (CI) = 0.48–0.59) among the expert panel, while for self-help group members $\kappa=0.23$ (95% CI = 0.19–0.27) [14]. The weighted κ was calculated by generating a κ score for each possible pair of raters for each item being rated. An overall κ score was then generated by calculating the average of individual κ with an appropriate overall standard error. The present interrater reliability findings for the DISCERN are also consistent with those reported for other studies [17].

Overview of DISCERN and BWQC ratings

Overall the quality of websites on bipolar disorder was disappointing. The mean total score on the DISCERN was 50.2 (SD = 7.6; possible range = 15–75), reflecting 58.7% of the maximum possible score. The mean total score for the BWQC was 180.3 (SD = 34.6, possible range = 56–280), reflecting 55.5% of the maximum possible score. The highest scoring subscales were Currency (mean = 8.7, SD = 2.4, possible range = 2–10), reflecting 83.8% of maximum possible score and Availability and Usability score, (mean = 17.1, SD = 2.1, possible range = 4–20), reflecting 81.9% of maximum possible score. The lowest scoring subscale was Breadth and Accuracy (mean = 105.3, SD = 30.3, possible range = 35–175), reflecting 50.2% of maximum possible score and Credibility (mean = 21.5, SD = 6.9, possible range = 7–35), reflecting 51.8% of the maximum possible score.

Comparison of DISCERN and BWQC against the site characteristics of Griffiths and Christensen

Table 3 details the mean scores for each subscale of the BWQC and the total score for the DISCERN grouped by the site characteristics of Griffiths and Christensen [12], that is, site ownership type, presence of editorial board, and scope of information. There was a significant difference in Credibility scores between sites that had an editorial board (mean = 27.1, SD = 3.7)

and sites that did not (mean = 16.7, SD = 5.0; $t = -4.6$, $df = 13$, $p = 0.001$). Websites with an editorial board were more likely to include details on credentials of authors and to provide references for statements made on the site. They were also more likely to subscribe to a voluntary code of ethics/conduct and to provide contact details. A significant difference was initially found for Mean Credibility scores with bipolar-plus websites and websites sites that contained bipolar info only (mean = 24.1, SD = 6.2 vs mean = 16.4, SD = 5.6, $t = -2.4$, $df = 13$, $p = 0.034$). This did not remain significant after applying the Bonferroni correction for multiple testing. Again a significant difference in the Objectivity scores between websites that were bipolar plus (bipolar plus: mean = 22.7, SD = 3.8) and sites with bipolar info only (mean = 17.9, SD = 4.2; $p = 0.045$) was not sustained after correction for multiple testing.

A series of one-way between-groups analysis of variance (ANOVA) was conducted to explore the impact of ownership type (professional, commercial or patient) on the mean scores for subscales of the BWQC, the total BWQC score and the total DISCERN score. There was initially a statistically significance difference between groups at the $p < 0.05$ level for the subscale Mean Credibility, ($F = 4.1$, $df = 2$, $p = 0.04$), but this did not remain significant after applying Bonferroni correction. Scheffé post-hoc tests were still performed that found that this initial significant difference was due to the professional websites' Mean Credibility scores (mean = 29.8, SD = 2.8) being greater than the Mean Credibility score for Commercial websites (mean = 18.8, SD = 6.9; $p = 0.044$). The patient websites' Mean Credibility scores (mean = 21.6, SD = 1.2) did not differ significantly from either the professional or commercial websites.

Factorial structure of the BWQC

To explore the structure of the BWQC, a principal components analysis (PCA) was performed. To ensure that the data set was suitable for conducting such an analysis, the sample size and strength of the relationship among the variables were considered. The Kaiser–Meyer–Oklin value was 0.68, reaching the recommended value of 0.6 and thereby supporting the factorability of the correlation matrix. The PCA showed two components with eigenvalues exceeding 1, explaining 45.6% and 20.7% of the variance, respectively. Inspection of the scree plot supported retaining the two components model for further investigation. Component 1 (consisting of four subscales with high loadings: Credibility, Objectivity, Currency and Breadth and Accuracy, i.e. 'Substance and Detail') contributed 46.6% of the variance, and component 2 (consisting of two subscales: Mean Availability and Usability and Mean Design and Aesthetics, i.e. 'Usability and Accessibility') contributed 20.7% based on the loading before rotation.

To aid the interpretation of these components, varimax rotation was performed (Table 4). The data from the BWQC were then re-analysed, collapsing the six original subscales into the two components, and their means compared using ANOVA with Games–Howell post-hoc analysis (selected due to the unequal sample sizes) against the three main website characteristics (Table 5). There was a significant relationship between scores on

Table 3. Website characteristics and scores for BWQC and DISCERN (mean ± SD)

Site characteristic	BWQC Subscales					Total	Mean	
	Credibility (Max = 35)	Currency (Max = 10)	Objectivity (Max = 30)	Availability and Usability (Max = 20)	Design and Aesthetics (Max = 10)			Breadth and Accuracy (Max = 175)
Ownership type								
Professional (n = 3)	29.8 ± 2.8 [§]	10.0 ± 0.0	24.8 ± 2.8	18.9 ± 0.4	8.5 ± 1.4	102.3 ± 51.3	194.2 ± 55.8	55.6 ± 9.9
Commercial (n = 9)	18.8 ± 6.9 [§]	8.0 ± 2.9	18.9 ± 4.0	16.8 ± 2.0	6.1 ± 1.7	103.5 ± 24.2	172.1 ± 30.2	47.8 ± 7.5
Patient (n = 3)	21.6 ± 1.2	9.3 ± 0.7	24.0 ± 3.2	16.3 ± 3.0	5.8 ± 1.9	113.9 ± 36.4	190.9 ± 40.8	52.1 ± 2.9
Editorial Board								
Yes (n = 7)	27.1 ± 3.7 [†]	9.8 ± 0.4	23.2 ± 2.6	17.1 ± 2.3	7.3 ± 1.6	102.9 ± 32.9	187.5 ± 33.2	52.5 ± 7.8
No (n = 8)	16.7 ± 5.0 [†]	7.7 ± 2.9	19.3 ± 5.1	17.1 ± 2.1	5.8 ± 1.8	107.4 ± 30.1	174 ± 37.4	48.2 ± 7.4
Scope of Information								
Bipolar info only (n = 5)	16.4 ± 5.6 [‡]	8.0 ± 2.2	17.9 ± 4.2 [‡]	16.8 ± 2.0	6.1 ± 1.8	101.1 ± 33.6	166.3 ± 35.4	47.5 ± 8.7
Bipolar plus (n = 10)	24.1 ± 6.2 [‡]	9.0 ± 2.5	22.7 ± 3.8 [‡]	17.3 ± 2.2	6.7 ± 1.9	107.4 ± 30.2	187.3 ± 36.4	51.6 ± 7.1

BWQC, Bipolar Website Quality Checklist. [†]Significant results (p < 0.006) after Bonferroni correction; [‡]significant results at p < 0.05; [§]significant after Scheffé post hoc.

Table 4. Principal component analysis with varimax rotation

BWQC subscales	Component	
	1: Substance/Detail	2: Usability/Accessibility
Credibility	0.895	
Objectivity	0.865	
Currency	0.803	
Breadth and Accuracy	0.411	
Availability and Usability		0.918
Design and Aesthetics	0.368	0.783

BWQC, Bipolar Website Quality Checklist.

component 1 (Substance and Detail) and ownership type (F(2,12) = 2.5, p = 0.001). The mean score for professional websites (mean = 200.4, SD = 18.2) was significantly greater than that for the patient websites (mean = 111.4, SD = 14.1), which in turn was significantly less than those for commercial websites (mean = 157.1, SD = 24.1). Although the professional website mean score was greater than that of the commercial websites this was not statistically significant. There was also a significant difference between the mean scores on component 2 (Usability and Accessibility) for those websites that had an editorial board (mean = 25.5, SD = 2.6) and those that did not (mean = 22.0, SD = 3.3; t(13) = -2.3, p = 0.04). That is, websites with an editorial board were more likely to present content in a clear and easily readable manner. Such websites used images that enhanced usability, and were more likely to have simple navigation tools and functional links to other sites. They were also more likely to provide easy access to information on bipolar disorder from the home page.

Correlations between the DISCERN and BWQC instruments

There was a strong positive correlation between the mean (calculated from the three raters), total DISCERN score (calculated by summing questions 1–15), and total mean BWQC scores (r = 0.78, p = 0.001). There was also a strong positive correlation between the mean BWQC component 1 (Substance and Detail) score and the mean total DISCERN score (r = 0.78, p = 0.001). There was no significant correlation between DISCERN and BWQC component 2 (Usability and Accessibility).

Ranking of websites

A major aim of the present study was to produce a rank order of the quality of websites to provide a guide for clinicians and patients. Table 6 details such ranking by the mean total scores on the DISCERN and BWQC. Four of the top six websites overlapped between these two measures, as did four of the lowest five.

The highest ranking website with both instruments was www.blackdoginstitute.org.au, the website for the Black Dog Institute in Sydney, Australia. The biggest discrepancy between the two lists was the ranking of the US National Institute of Mental Health (NIMH) website (www.nimh.nih.gov). The NIMH website was ranked second on the DISCERN and seventh on the BWQC. Another website that differed substantially in its relative rankings was the patient website www.helpguide.org, which was ranked ninth on DISCERN and second on the BWQC.

In order to further examine these discrepancies in the rankings of the NIMH and Black Dog Institute websites, the mean scores for the BWQC were examined at an individual item level. Both websites had very similar scores on the BWQC subscales Currency, Availability/Usability and Design and Aesthetics. On the Credibility subscale, the Black Dog website scored more highly on the item 'code of conduct' (5 vs 2.3) and for 'presence of a quality marker' on the home page (5 vs 1). The NIMH website scored more highly on items 'credentials shown for authors' (5 vs 3) and 'references shown for content' (5 vs 4.3). The greatest difference, however, between these two websites was within the subscale Breadth/Accuracy. Although both websites scored highly on items focusing on signs and symptoms and aetiology of bipolar disorder as well as the item on the risk and management of suicide, the Black Dog Institute website overall scored more highly on the majority of the other items because it provided more detailed information on a variety of treatments both biological and psychological and covered in more detail issues around differential diagnosis, comorbid conditions and hospitalization. For example, mean scores for the items 'difference in treatment for bipolar I vs bipolar II disorder' (5 vs 2.3), 'combination of mood stabilizers' (4 vs 2.7), 'role of atypical anti-psychotics' (4 vs 2.3), and 'role of emerging treatments e.g. Omega 3' (5 vs 3). The Black Dog Institute also scored more highly on items on psychoeducation (5 vs 3.7) and role of lifestyle in keeping well (4 vs 2.3). The BWQC was able to evaluate the quality of the

content of the website, which the DISCERN is not able to do as a generic quality rating instrument.

Discussion

This paper has detailed the development of the first website quality evaluation instrument specifically designed to assess bipolar disorder websites. This instrument, the BWQC, consists of 56 items and six subscales (Credibility; Currency; Objectivity; Availability and Usability; Design and Aesthetics; and Breadth and Accuracy), which were derived from previously developed concepts on the appropriate means for measuring website and print media quality [10–12]. PCA demonstrated a structure of two main components (Substance and Detail, and Usability and Accessibility), which together accounted for 67% of the variance. The BWQC demonstrated high interrater reliability ($r = 0.89$). The BWQC correlated strongly ($r = 0.78$) with the more generic DISCERN instrument [14].

Although the top-ranked website in terms of hits with common search engines was the US patient site www.helpguide.org, the leading website in terms of quality (as rated by both the BWQC and the DISCERN) was that of the Black Dog Institute in Sydney, Australia (www.blackdoginstitute.org.au).

Overall, there were only minor variations in the ranking of the websites by the two instruments. The exceptions were the discrepant rankings of the NIMH website (www.nimh.nih.gov) and the aforementioned patient site www.helpguide.org. The NIMH site was ranked second of 15 on the DISCERN, but seventh on the BWQC. The probable reason for this low ranking was that this website scored poorly on the subscale Breadth and Accuracy, reflecting its limited amount of detailed information. This particular subscale was based on the guideline score of Griffiths and Christensen [12], with the detail being derived from the recommendations of the Royal Australian and New Zealand College of Psychiatrists' clinical guidelines for bipolar disorder [7]. The specific details of these guidelines were very similar to those of the American Psychiatric Association [8].

The present search strategy found few websites linked to government mental health service providers or educational institutions. The majority of websites were affiliated to private organizations or commercial entities. Although major search engines do not make the details on how they rank websites publicly available, it is acknowledged that webmasters and web developers are able to manipulate how they

Table 5. Website characteristics scores on the two components of the BWQC

Site characteristic	Component 1 'Substance and Detail' (Mean ± SD)	Component 2 'Usability and Accessibility' (Mean ± SD)
Ownership type		
Professional (n=3)	200.4 ± 18.2*	27.0 ± 2.2
Commercial (n=9)	157.1 ± 24.1*	22.2 ± 3.1
Patient (n=3)	111.4 ± 14.1*	24.6 ± 3.2
Editorial board		
Yes (n=7)	169.0 ± 33.6	25.5 ± 2.6*
No (n=8)	145.9 ± 35.5	22.0 ± 3.3*
Scope of information		
Bipolar info only (n=5)	150.7 ± 37.2	22.0 ± 3.4
Bipolar plus (n=10)	159.6 ± 36.2	24.5 ± 3.3

BWQC, Bipolar Website Quality Checklist. *p < 0.05.

present their websites to a search engine in order to improve their numerical ranking. Google's popularity as a search engine has in part been due to its development of the algorithm PageRank, which has made its ranking system more resistant to such manipulation. As well as examining keywords within the text, it also assesses the page's strength and quality of inbound links. It may be that other high-quality bipolar disorder websites developed by government agencies and tertiary educational facilities were not included in this survey due to low ranking on popular search engines such as Yahoo and Google.

It should be noted, however, that recent work by Griffiths and Christensen reported that evidence-based quality scores correlated significantly with Google page rank [18]. Those authors suggest that this may indicate that Google page rank had 'promise as an automatic indicator of quality' and was moving towards being an acceptable marker of quality. Although this finding may reflect slowness on the part of professional institutions in adopting the Internet as a medium of disseminating information, it may also reflect that commercial websites have been developed with a greater emphasis on gaining a high website ranking on these search engines. Organizations and institutions producing information on bipolar disorder for the Internet should therefore be aware of how search engines identify and rank sites.

Limitations

A major limitation of this survey was the lack of patient involvement in the ranking process. The importance of this has been raised by Eysenbach *et al.* [10], but recent work by Griffiths and Christensen in their cross-sectional survey of quality of depression websites using both health professional and patient raters found a significant positive correlation between patient and health professional ratings using the DISCERN ratings [18].

Two out of the three raters had affiliations with the organization that created the top-ranked website www.blackdoginstitute.org.au. It is possible that, despite the objective nature of the two instruments, there could have been a halo effect despite the intent to remain impartial. It should be noted that the third independent rater had no affiliation with the Black Dog Institution. All reviewers were Australian, which, it could be argued, may also have had an impact on this result. It should be noted, however, that six out of the seven search engines were international, making it unlikely that the initial selection of websites for review based on the ranking by these search engines was biased towards an Australian website.

Because a major subscale within the BWQC (Breadth and Accuracy) was developed from clinical treatment guidelines, it possible that it had an inherent bias towards rating diagnostic and clinical information more highly than other types of information. This may

Table 6. Website ranking vs search engine, mean total DISCERN and BWQC

Website	Ranking based on DISCERN	Mean DISCERN score	Ranking based on BWQC	Total mean BWQC score	Website type
blackdoginstitute.org.au	1	64	1	248	1
nimh.nih.gov	2	58	7	198	1
bipolar.about.com	3	56	3	207	2
bipolarworld.net	4	55	5	199	3
medicine.net.com	5	54	4	204	2
netdoctor.co.uk	6	53	5	199	2
en.wikipedia.org	6	53	8	195	2
patient.co.uk	6	53	9	165	2
helpguide.org	9	52	2	227	3
pendulum.org	10	49	12	147	3
nlm.nih.gov	11	45	14	137	1
answers.com	12	43	13	139	2
bipolar.com	13	41	15	125	2
mentalhealthchannel.net	13	41	10	162	2
manic-depression.net	15	36	11	154	2

BWQC, Bipolar Website Quality Checklist. Ownership type: 1 =professional, 2 =commercial, 3 =patient.

make it more difficult for patient websites to obtain high scores on this particular subscale.

Finally, although over 140 websites were initially identified using the search strategy described in this paper, only 15 were finally reviewed in detail. This relatively small sample size may limit generalization of some of the findings, especially when distinguishing between professional, commercial and patient websites. Clearly, further research by other groups will be necessary to confirm or refute these findings. The authors would recommend that any future reviews attempt to sample a larger sample of websites and include a cross-section of potential end-users such as patients and carers, as raters.

Conclusions

Overall, the quality of information on bipolar disorder on the Internet appears to be variable, which is a somewhat different conclusion from an earlier review that reported the quality to be good [19]. Those authors found a negative relationship between readability and interactivity of a website and its content quality, and commented that the DISCERN score could not predict the quality of content of a website. This finding has been noted by the developers of the DISCERN [14] and is further supported by the present findings. It appears that there are some excellent websites that are easily accessible that not only provide high-quality detailed information on bipolar disorder and its treatment, but furthermore are using the Internet's unique interactive elements to provide many other useful features. The low ranking (14th), however, by the Internet search engines of www.blackdoginstitute.org.au (the highest quality ranking website, according to our evaluation), is of concern because this makes it less likely that this website would be found by a casual browser. When developing websites, search engine optimization (e.g. through meta tags and key word attribution or descriptors) will be crucial if websites are to be easily accessed.

Because this review was undertaken in August 2007, a repeat of the website search strategy today would be likely to identify new websites, and to find some of those identified in this paper to have been markedly modified. This fluidity is one of the Internet's most striking features. In reading this review clinicians cannot solely rely on its 2007 findings, but are encouraged to reflect on these when seeking out and recommending contemporary websites to patients. The authors also believe that it is

important that clinicians attempt to match websites with individual needs, because patient websites often have many useful interactive features such as chat rooms and discussion forums that provide support and promote self-help and self-management. These features could not be rated with the BWQC or DISCERN instruments, and are often missing from commercial websites.

The Internet is increasingly becoming the portal of choice for patients and carers in accessing information on health issues. Clinicians therefore need to be able to assist patients to avoid the 'bad and the ugly' [19] and should be able to recommend quality websites as part of their routine provision of information on bipolar disorder. By doing so, clinicians are meeting current treatment guidelines that highlight the importance of psychoeducation as a crucial component of the management of bipolar disorder.

Acknowledgements

The research activities of Professor Philip Mitchell and Dusan Hadzi-Pavlovic are supported by an Australian National Health and Medical Research Council Program Grant (510135). Dr Caryl Barnes' research was supported by a doctoral scholarship grant through the School of Psychiatry, University of New South Wales, Sydney. Ms Wilde and Dr Barnes have an affiliation with the Black Dog Institute as doctoral students at the University of New South Wales.

References

1. Berland G, Elliott M, Morales L. Health Information on the Internet: accessibility, quality and readability in English and Spanish. *JAMA* 2001; 285:2612-2621.
2. MiniWatts Marketing Group. *Internet World Statistics*. Colombia: MiniWatts Marketing Group, 2008.
3. Fox S. *Health information online*. Washington, DC: Pew Internet and American Life Project, 2005.
4. Fox S, Rainie L. *The online health care revolution: how the web helps Americans take better care of themselves*. Washington, DC: Pew Charitable Trusts, 2000.
5. Colom F, Lam D. Psychoeducation: improving outcomes in bipolar disorder. *Eur Psychiatry* 2005; 20:359-364.
6. Colom F, Vieta E, Reinares M *et al*. Psychoeducation efficacy in bipolar disorders: beyond compliance enhancement. *J Clin Psychiatry* 2003; 64:1101-1105.
7. Mitchell PB, Malhi GS, Redwood BL, Ball JF, RANZCP Clinical Practice Guideline team for Bipolar Disorder. Summary of the guidelines for the treatment of bipolar disorder. *Australas Psychiatry* 2003; 11:39-53.
8. American Psychiatric Association. Practice guidelines for treatment of bipolar disorder. *Am J Psychiatry* 2002; 159: 1-50.

9. Hirschfeld RA, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive Association 2000 survey of individuals with bipolar disorder. *J Clin Psychiatry* 2003; 64:161–174.
10. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* 2002; 287:2691–2700.
11. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling and assuring quality of medical information on the internet. *JAMA* 1997; 277:1244–1245.
12. Griffiths K, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. *BMJ* 2000; 321:1511–1515.
13. Morel V, Chatton A, Cochand S, Zullino D, Khazaal Y. Quality of web-based information on bipolar disorder. *J Affect Disord* 2008; 110:265–269.
14. 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; 53:105–111.
15. Christensen H, Griffiths KM, Medway J. Sites for depression on the web; a comparison of consumer, professional and commercial sites. *Aust N Z J Public Health* 2000; 24:396–400.
16. Depression Guideline Panel. *Depression in Primary Care. Clinical Practice Guidelines No. 5*. Bethesda: US Department of Health and Human Services, 1993.
17. Rees CE, Ford JE, Sheard CE. Evaluating the reliability of DISCERN: a tool for assessing the quality of written patient information on treatment choices. *Patient Educ Couns* 2002; 47:273–275.
18. Griffiths KM, Christensen H. Website quality indicators for consumers. *J Med Internet Res* 2005; 7:e55.
19. Wilson P. How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the internet. *BMJ* 2002; 324:598–600.

Appendix 1. Bipolar Quality Website Checklist

URL: _____ *Date:* _____ *Reviewer:* _____

Quality Criteria	Rating Scale				
	1 No	2	3 Partially	4	5 Yes
1. Credibility					
1.1	Does the URL indicate reputable affiliation?				
1.2	Is there a quality marker associated with the website?				
1.3	Does the website show affiliation to code of conduct?				
1.4	Is there contact information available for website?				
1.5	Is there an overall editorial/ review process in place?				
1.6	Are sources of information/ references shown?				
1.7	Are the credentials of authors shown?				
2. Currency					
2.1	Is the date of website creation shown?				
2.2	Are the pages stamped with date of last update?				
3. Objectivity					
3.1	Are aims of the site clearly stated?				
3.2	Is the intended audience made clear?				
3.3	Is there disclosure of sponsorship/ conflict of interest?				
3.4	Is there advertising on the same pages as factual information is presented?				
3.5	Is the information presented as factual?				
3.6	Are areas of debate/uncertainty discussed?				
4. Availability and Usability					
4.1	Site contains current external links to other sites?				
4.2	Search tools present in website				
4.3	Navigation tools/internal links allow easy access around site/ return to homepage				
4.4	Information on bipolar disorder readily found from homepage				
5. Design and Aesthetics					
5.1	Text clear and presented in easy to read sections				
5.2	Images used in a manner that enhances use of site				
6. Breadth and Accuracy					
6.1	Accurate information presented on clinical characteristics of disorder (subtypes BPAD I, II and types of episodes described).				
6.2	Description of aetiology present and accurate				
6.3	Seriousness of disorder and risk of suicide discussed				
6.4	Biological and psychological treatments both described as being effective and important				
6.5	Treatment differences between bipolar I and bipolar II discussed				
6.6	Lithium 'bench mark/gold standard'				
6.7	Carbamazepine				
6.8	Sodium Valproate				
6.9	Lamotrigine				
6.10	Role of atypical anti-psychotics as mood stabilizers				
6.11	ECT				
6.12	Combination treatment (two mood stabilizers)				
6.13	Combination treatment (mood stabilizers and antipsychotic)				
6.14	SSRIs				
6.15	TCAs-with mention of the increased risk of switching				
6.16	MAOI				
6.17	SNRIs e.g. Venlafaxine				
6.18	Role of Benzodiazepines				
6.19	Mention of new treatment developments e.g. omega 3				
6.20	Role of typical antipsychotics in mania				
6.21	Risks/ side-effects of each medication mentioned on site discussed				
6.22	Different types of treatment regimes for acute and maintenance treatment discussed				
6.23	Importance of maintenance treatment mentioned				
6.24	Drug levels/blood tests when on mood stabilizers mentioned				
6.25	Role of lifestyle changes in staying well				
6.26	Possible differential diagnosis to BPAD e.g. ADHD, BPD				
6.27	Common co-morbidities, importance of diagnosis and treatment				
6.28	Psycho-education				

Appendix 1. *Continued*

Quality Criteria	1	2	Rating Scale	4	5
	No		3 Partially		Yes
6.29	Cognitive therapy				
6.30	IPSRT				
6.31	Group therapy				
6.32	Family Focused Therapy				
6.33	Indications for inpatient treatment/hospitalization discussed				
6.34	Issues in treatment of BPAD in children/adolescents				
6.35	Issues in treatment of BPAD in pregnancy, breast feeding discussed				