
UNIVERSITE DE LAUSANNE - FACULTE DE BIOLOGIE ET DE
MEDECINE

Département Universitaire de Psychiatrie Adulte (DUPA)

QUALITY OF WEB-BASED INFORMATION ON BIPOLAR
DISORDER

THESE

préparée sous la direction du Professeur honoraire François Borgeat,
avec la collaboration du Docteur Yasser Khazaal, médecin adjoint du service
d'addictologie du Département de psychiatrie de l'Université de Genève

et présentée à la Faculté de biologie et de médecine de l'Université de Lausanne
pour l'obtention du grade de

DOCTEUR EN MEDECINE

par

Vincent MOREL

BMEC 3494

VM
207
Mor

Médecin diplômé de la Confédération Suisse

Originaire de Marnand (VD)

Lausanne

2008

Rapport de synthèse

Qualité de l'information sur Internet concernant le trouble bipolaire

Introduction : Internet est une source importante d'information sur la santé mentale. Le trouble bipolaire est communément associé à un handicap, des comorbidités, un faible taux d'introspection et une mauvaise compliance au traitement. Le fardeau de la maladie, de par les épisodes dépressifs et maniaques, peut conduire les personnes (dont le diagnostic de trouble bipolaire a été déjà posé ou non), ainsi que leur famille à rechercher des informations sur Internet. De ce fait, il est important que les sites Web traitant du sujet contiennent de l'information de haute qualité, basée sur les évidences scientifiques.

Objectif : évaluer la qualité des informations consultables sur Internet au sujet du trouble bipolaire et identifier des indicateurs de qualité.

Méthode: deux mots-clés : « bipolar disorder » et « manic depressive illness » ont été introduits dans les moteurs de recherche les plus souvent utilisés sur Internet. Les sites Internet ont été évalués avec un formulaire standard conçu pour noter les sites sur la base de l'auteur (privé, université, entreprise,...), la présentation, l'interactivité, la lisibilité et la qualité du contenu. Le label de qualité « Health On the Net » (HON), et l'outil DISCERN ont été utilisés pour vérifier leur efficacité comme indicateurs de la qualité.

Résultats: sur les 80 sites identifiés, 34 ont été inclus. Sur la base de la mesure des résultats, la qualité du contenu des sites s'est avérée être bonne. La qualité du contenu des sites Web qui traitent du trouble bipolaire est expliquée de manière significative par la lisibilité, la responsabilité et l'interactivité aussi bien que par un score global.

Conclusions: dans l'ensemble, la qualité du contenu de l'étude des sites Web traitant du trouble bipolaire est de bonne qualité.



ELSEVIER

Journal of Affective Disorders xx (2008) xxx–xxx

JOURNAL OF
AFFECTIVE
DISORDERS

www.elsevier.com/locate/jad

Brief report

Quality of web-based information on bipolar disorder

Vincent Morel^b, Anne Chatton^a, Sophie Cochand^a, Daniele Zullino^a, Yasser Khazaal^{a,*}

^a Division of Substance Abuse, Geneva University Hospitals, Geneva, Switzerland

^b Department of Psychiatry, University Hospital of Lausanne, Switzerland

Received 10 September 2007; received in revised form 7 January 2008; accepted 7 January 2008

Abstract

Objective: To evaluate web-based information on bipolar disorder and to assess particular content quality indicators.

Methods: Two keywords, “bipolar disorder” and “manic depressive illness” were entered into popular World Wide Web search engines. Websites were assessed with a standardized proforma designed to rate sites on the basis of accountability, presentation, interactivity, readability and content quality. “Health on the Net” (HON) quality label, and DISCERN scale scores were used to verify their efficiency as quality indicators.

Results: Of the 80 websites identified, 34 were included. Based on outcome measures, the content quality of the sites turned-out to be good. Content quality of web sites dealing with bipolar disorder is significantly explained by readability, accountability and interactivity as well as a global score.

Conclusions: The overall content quality of the studied bipolar disorder websites is good.

© 2008 Elsevier B.V. All rights reserved.

Keywords: Internet; Quality indicators; Health care; Bipolar disorder

1. Introduction

Internet is a source of information on mental health issues (Powell and Clarke, 2006). Bipolar disorder (BPD) is commonly associated with disability (Judd and Akiskal, 2003), comorbidity (Yerevanian et al., 2001), low level of insight and lack of treatment adherence (Greenhouse et al., 2000; Scott and Pope, 2002). The burden of the disorder (in its depressive and manic episodes) may lead people (undiagnosed and diagnosed persons and their relatives)

to search on-line evidence-based information. This makes that it is important for websites to present high content quality (evidence-based-health information) on the topic.

It was repeatedly reported that content quality of websites is a problem (Eysenbach et al., 2002).

The present study aimed, firstly, to assess content quality and general quality (presentation features) of BPD websites. Secondly, the study aimed to determine content quality indicators.

2. Methods

A web search was performed using search engines available for the general user of Internet using widely used terms for the condition.

* Corresponding author. Division of Substance Abuse, Geneva University Hospitals, Rue verte, 2, 1205 Geneva, Switzerland. Tel.: +41 223725550; fax: +41 223202840.

E-mail address: yasser.khazaal@hcuge.ch (Y. Khazaal).

2.1. Selection of websites

Keyword searches and websites evaluation were done in November 2006 (Switzerland) by the coauthor (SC, psychologist). The keywords “bipolar disorder” and “manic-depressive illness” were entered into two search engines: Google and Yahoo.

The first 20 English language websites coming-up from each keyword query were examined, as most people rarely search beyond the first 20 retrieved links (Eysenbach et al., 2002).

Sites were excluded if: inaccessible (invalid address), already reviewed in the current study, containing no information on BPD, requiring access fee, discussion group or open forum, not a site (external links, books or articles) and no information in English. Exclusion of sites that were exclusively discussion group and forums was made considering that the method used in the present study is more adapted for evaluation of informative websites.

2.2. Evaluation of the websites

Websites affiliations were divided into 5 categories: commercial, university, non-profit organization, governmental, or other according to the suffix (i.e. .gov: government) and the declaration of affiliation (in order to explore potential links between affiliation and content quality). Presence of the Health On the Net foundation logo (HON) as an indicator of adhesion to some ethical standards (Boyer et al., 1998) was also considered.

Websites were assessed with a standardized proforma designed to rate sites on the basis of accountability, presentation, interactivity, readability and content quality. The proforma, based on previous studies (Abbott, 2000; Griffiths and Christensen, 2000; Khazaal et al., 2007; Kisely et al., 2003; Nilsson-Ihrfelt et al., 2004; Silberg et al., 1997) was adapted in order to avoid overlap between instruments (i.e between Silberg and Abbott scales). The proforma (Table 1) assessed accountability (Silberg scale) (Silberg et al., 1997), interactivity (Abbott, 2000), Abbott's aesthetic criteria (Kisely et al., 2003). The full Abbott's scale was not used due to partial overlap with Silberg score. Interactivity and aesthetic issues were adapted from the Abbott's scale (Khazaal et al., 2007).

Content quality was evaluated on the availability of responses to probable queries. Questions focused on advice for treatment and information concerning BPD, as it was previously shown that these kind of questions are the most common patients and relative queries on the Internet (Shuyler and Knight, 2003). The retrieved

Table 1
Components of the proforma

	Scoring method	
	Present=1	Absent=0
Silberg scale (accountability)		
Authors credited		
Authors affiliations		
Authors credentials		
Information sources given		
References given/hyperlinked		
Site ownership disclosed		
Sponsorship disclosed		
Site modified in the last 6 months		
Date created/last modified specified		
Interactivity		
Intra-site search engine		
Audio or video support		
Evaluation questionnaires		
Supporting bodies (forums, discussion rooms)		
Possibility to send queries to a webmaster		
Abbott aesthetic criteria		
Headings		
Diagrams		
Hyperlinks to external sites		
Absence of advertisement		
Content quality	Coverage	Correctness
	None=0	Mostly not=0
	Minimal=1	Mostly=1
	Sufficient=2	Completely right=2
« How can I know if I am suffering from BPD? » (definition)		
« What is the origin of this disorder? If have this disorder, what will happen to my children? » (etiology)		
« What are the most efficacious treatments? » (treatment availability)		
« If I have to take a mood stabilizing agent, for how long will I need to take it? And what are the common side effects? » (mood stabilizing agents)		
« Can I take an antidepressant agent? For how long do I need to take it, and for what? » (antidepressant drugs)		
« Can I take an antipsychotic agent? For how long do I need to take it and for what? » (antipsychotic drugs)		
« What forms of psychotherapy are effective for the treatment of BPD? Can I undergo psychotherapy with or without a pharmacological treatment? » (psychotherapy).		
Readability *		
Flesch–Kincaid grade level		
Flesch–Kincaid readability index		

* Readability scores were calculated with mathematical formulas treating number of words and sentences.

information was compared to official guidelines (American Psychiatric Association, 2006: Practice guidelines for the treatment of BPDs). Content quality score is the sum of coverage (exhaustibility) and accuracy (correctness) (Nilsson-Ihrfelt et al., 2004). Similarly than previously described (Griffiths and Christensen, 2000) (Khazaal et al., 2007), a global score was defined as the sum of Silberg, interactivity, adapted Abbott's aesthetic criteria and content quality.

Readability was assessed using the Flesch–Kincaid grade level score and the Flesch–Kincaid readability index (Kisely et al., 2003). The first score evaluates the degree of text reading difficulty with regard to USA school grades. Eight, the recommended level for standard documents, means that an eighth grade student can easily understand the document. The second score is included in the Microsoft word spellchecker and ranges from 0 to 100, with higher scores reflecting higher legibility.

Finally, the DISCERN was used. This instrument intends to assist people without content expertise to assess written health attributes of a publication (with a particular focus on the content of the information on treatment choices), such as the extent to which the information appears unbiased (Charnock et al., 1999; Griffiths and Christensen, 2002). The instrument comprises 16 items each rated from 1 to 5. An association was previously found between adherence to clinical practice guidelines and DISCERN (Griffiths and Christensen, 2002). Previous studies showed that consumers and health professional DISCERN's ratings were significantly correlated (Griffiths and Christensen, 2005). Inter-rater reliability of proforma scores was assessed based on a random sample of 15 sites with 2 trained evaluators (a medical doctor VM and SC).

2.3. Analyses

Statistical analyses were performed using SPSS for Windows (version 11.0). Student's T-tests were used to compare sites having the HON label with sites without this label.

The main analysis consisted in the prediction of the content quality of websites through a multiple linear regression. Still, before doing so, factor analysis was first used to reduce the number of interrelated variables with quantitative values to a more meaningful, smaller set of uncorrelated variables. Five original variables were reduced: Silberg, Interactivity, Abbott aesthetic criteria scores, Flesch reading ease and the Flesch–Kincaid education scores. The Global score and DISCERN were not considered in this Factor analysis since they are two

different tools used to evaluate the content and general quality of a site. To test whether Factor analysis was satisfactory, we used the Kaiser–Meyer–Olkin (KMO) and Bartlett's tests.

Two criteria for deciding how many factors to retain were observed: the Kaiser rule (factors with Eigenvalues greater than 1) and the Cattell scree plot. The extraction of the factors was made by variance maximizing (varimax) rotation of the original variable space, ensuring that these factors are uncorrelated or orthogonal to each other.

Finally, the extracted significant factors, global score, DISCERN and categorical variables were analyzed through a multiple linear regression for the prediction of content quality. For all analyses, a significance level of $p \leq 0.05$ was used.

3. Results

We reviewed 80 websites. There was an overlap among the sites identified by the search engines and the 2 keywords (35/80: 43.8%). This left 45 websites (80–35) from which 11 were excluded for these reasons: one required an access fee; 8 were not websites (only external links or books), 2 were not in English. We therefore included 34 websites (list available upon request).

Websites were assessed with the proforma which components resulted to have a good inter-rater reliability: Silberg ($r=0.841$; $p<0.05$), Flesch Reading ease ($r=0.881$; $p<0.05$), Flesch Kincaid level ($r=0.835$; $p<0.01$), Abbott's aesthetic criteria ($r=0.751$; $p<0.05$), DISCERN ($r=0.942$; $p<0.01$), content quality ($r=0.851$; $p<0.01$), interactivity ($r=0.865$; $p<0.01$).

Origin of sites was as follows: (government: 1, university: 4, commercial: 15, non-profit organization: 12, others, i.e. personal pages, blogs: 2). Among the 34 sites, 15 (44%) had the HON logo. Twenty-three percent of sites presented additional information in Spanish and 9% in other languages. Intra-site search engine appeared in 17.6% of sites, a support group in 20.6% and the possibility to send queries to a web master in 11.8%. Most sites specified the date of launch (61.8%), 29.4% a date of modification during the last six months. Content quality (sum of total coverage and total accuracy) mean score was 19 out of 28 and was highly variable (6 to 27).

Means and standard deviation scores are described in Table 2. Sites holding the HON label scored lower than sites without this label on Flesch–Kincaid education scores ($t=2.39$; $p=0.02$). No other statistically significant differences between these groups were observed.

The KMO measure is above the recommended value (0.52) and the Bartlett's test is significant at $p<0.0005$.

Table 2
Websites' general and content quality indicators, mean±SD, (N=34)

Outcome	Mean (SD)
Silberg scores (0-9)	4.8 (2.2)
Interactivity scores (0-6)	1.8 (1.2)
Abbott aesthetic criteria scores(0-4)	2.8 (0.6)
Flesch reading ease scores (0-100)	30.1 (12.7)
Flesch-Kincaid education scores	10.3 (2)
Content quality scores (0-28)	19 (4.8)
Global scores (0-47)	28.3 (6)
DISCERN scores (16-80)	43.7 (11.9)

By Principal components method, 2 factors have been extracted. "Flesch reading ease and Flesch-Kincaid education scores" load on factor 1 labeled "readability". Silberg and "Interactivity" load on factor 2 named "accountability and interactivity". These 2 factors accounted for almost 68% of item variance, with factor 1 accounting for 42.34% and factor 2 for 25.32%.

These 2 factors, global score, site affiliation (commercial vs. non commercial) and HON label (yes vs. no) were analyzed through a multiple linear regression to predict content quality. Non commercial sites and sites with HON label were taken as the reference category. Preliminary partial correlation test shows no association between the dependent variable and the DISCERN variable when controlled for global score. Therefore, DISCERN was not considered in the regression.

The overall *F* statistic of the regression is significant at $p < 0.0005$. Adjusted R^2 statistic showed that almost 99% of the total variation in outcome was explained by a final model including: factor1 (readability), factor2 (accountability and interactivity), and global scores but not by site affiliation or HON label. An outlier had to be deleted.

The final model was: Content quality = $\beta_0 + \beta_1$ readability + β_2 accountability and interactivity + β_3 globalscore + ε (ε : error term).

β_0 , β_1 , and β_2 and their associated confidence interval are:

-9.2 [-9.9, -8.4], -0.4 [-0.5, -0.2], -2.8 [-3, -2.7], 1[0.97, 1.02] respectively. So the estimated model is:
Content quality = $-9.2 - 0.4 \times$ readability $- 2.8 \times$ accountability and interactivity + global score (predicted mean content quality: 19 ± 4.8)

Assumptions of independent normally-distributed errors and constant variance of errors were checked by studying the residuals from the model. Potential colinearity among the independent variables was studied through the variance inflation factor (VIF) value. Fi-

nally, goodness of fit of the model was assessed through the coefficient of determination: the adjusted R^2 statistic. No abnormal pattern was observed.

The results suggest that content quality of web sites dealing with BPD is significantly explained by readability, accountability and interactivity as well as global scores (all p -values < 0.0005) but not by presence of HON label and site origin.

4. Discussion

The present study aimed to assess Internet information on BPD and to determine content quality indicators. In contradiction with results of previous studies on health related web sites (Eysenbach et al., 2002; Griffiths and Christensen, 2000; Kisely et al., 2003; Wyatt, 1997), content quality of BPD related web sites could be considered as good (score 19/28: higher than 14, the mean score for a minimal and mostly right information).

The two factors issued from the factor analysis and the global score seem to predict respectively negatively and positively, the content quality. Astonishingly, the study found a negative relation between readability as well the factor 2 (accountability and interactivity) and content quality. It seems that most Internet sites choose either to promote the quality of their content, either to promote other aspects such as interactivity and rarely these two aspects concomitantly.

Furthermore, the DISCERN score did not predict content quality which also contrasts with some previous studies (Griffiths and Christensen, 2002). The present result was however found when controlling the association between DISCERN and content quality for global score.

HON label failed to predict content quality as well as websites affiliation (non commercial vs. commercial).

The present study has several limitations. It only gives account of the situation on BPD related websites in November 2006. Two general search engines were used rather than medical search engines which, are however more commonly used by general population, and lead to similar results as previously shown (Ilic et al., 2003). The study may not reflect behaviors of all Internet users who possibly use variable search methods in regard to search engines and keywords used which possibly vary for example between diagnosed and undiagnosed persons. Additionally, the study has not distinguished between websites aiming to give information and those who primarily intend to be interactive. Finally, some criteria such as Silberg score are issued from guiding principles rather than from a standardized validation.

Nonetheless, this study brings to evidence the good content quality of BPD related websites. A possible method to improve BPD websites would probably be to associate content quality with better readability, accountability and interactivity patterns or to promote two types of websites “Informative” vs “Interactive”.

Role of funding source
Nothing declared.

Conflict of interest
No conflict declared.

Acknowledgment

To Delphine Sreekumar for the revision of the paper syntax and grammar.

References

- Abbott, V.P., 2000. Web page quality: Can we measure it and what do we find? A report of exploratory findings. *Journal of Public Health Medicine* 22, 191–197.
- Boyer, C., Selby, M., Appel, R.D., 1998. The health on the net code of conduct for medical and health web sites. *Medinfo* 9 (Pt 2), 1163–1166.
- Charnock, D., Shepperd, S., Needham, G., Gann, R., 1999. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology and Community Health* 53, 105–111.
- 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, 2691–2700.
- Greenhouse, W.J., Meyer, B., Johnson, S.L., 2000. Coping and medication adherence in bipolar disorder. *Journal of Affective Disorders* 59, 237–241.
- Griffiths, K.M., Christensen, H., 2000. Quality of web based information on treatment of depression: Cross sectional survey. *BMJ. British Medical Journal (Clinical Research ed.)* 321, 1511–1515.
- Griffiths, K.M., Christensen, H., 2002. The quality and accessibility of Australian depression sites on the World Wide Web. *Medical Journal of Australia* 176, S97–S104 Suppl.
- Griffiths, K.M., Christensen, H., 2005. Website quality indicators for consumers. *Journal of Medical Internet Research* 7, e55.
- Ilic, D., Bessell, T.L., Silagy, C.A., Green, S., 2003. Specialized medical search-engines are no better than general search-engines in sourcing consumer information about androgen deficiency. *Human Reproduction (Oxford, England)* 18, 557–561.
- Judd, L.L., Akiskal, H.S., 2003. The prevalence and disability of bipolar spectrum disorders in the US population: Re-analysis of the ECA database taking into account subthreshold cases. *Journal of Affective Disorders* 73, 123–131.
- Khazaal, Y., Fernandez, S., Cochand, S., Reboh, I., Zullino, D., 2007. Quality of web-based information on social phobia: A cross-sectional study. *Depress Anxiety*.
- Kisely, S., Ong, G., Takyar, A., 2003. A survey of the quality of web based information on the treatment of schizophrenia and attention deficit hyperactivity disorder. *Australian and New Zealand Journal of Psychiatry* 37, 85–91.
- Nilsson-Ihrfelt, E., Fjallskog, M.L., Blomqvist, C., Ahlgren, J., Edlund, P., Hansen, J., Malmberg, L., Villman, K., Andersson, G., 2004. Breast cancer on the Internet: the quality of Swedish breast cancer websites. *Breast (Edinburgh, Scotland)* 13, 376–382.
- Powell, J., Clarke, A., 2006. Internet information-seeking in mental health: Population survey. *British Journal of Psychiatry* 189, 273–277.
- Scott, J., Pope, M., 2002. Self-reported adherence to treatment with mood stabilizers, plasma levels, and psychiatric hospitalization. *The American Journal of Psychiatry* 159, 1927–1929.
- Shuyler, K.S., Knight, K.M., 2003. What are patients seeking when they turn to the Internet? Qualitative content analysis of questions asked by visitors to an orthopaedics Web site. *Journal of Medical Internet Research* 5, e24.
- Silberg, W.M., Lundberg, G.D., Musacchio, R.A., 1997. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewer — Let the reader and viewer beware. *JAMA* 277, 1244–1245.
- Wyatt, J.C., 1997. Commentary: measuring quality and impact of the World Wide Web. *BMJ. British Medical Journal (Clinical Research ed.)* 314, 1879–1881.
- Yerevanian, B.I., Koek, R.J., Ramdev, S., 2001. Anxiety disorders comorbidity in mood disorder subgroups: data from a mood disorders clinic. *Journal of Affective Disorders* 67, 167–173.