HON Label and DISCERN as Content Quality Indicators of Health-Related Websites

Yasser Khazaal · Anne Chatton · Daniele Zullino · Riaz Khan

Abstract Content quality indicators are warranted in order to help patients and consumers to judge the content quality of health-related on-line information. The aim of the present study is to evaluate web-based information on health topics and to assess particular content quality indicators like HON (Health on the Net) and DISCERN. The present study is based on the analysis of data issued from six previous studies which assessed with a standardized tool the general and content quality (evidence-based health information) of health-related websites. Keywords related to Social phobia, bipolar disorders, pathological gambling as well as cannabis, alcohol and cocaine addiction 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 (evidence-based information). "Health on the Net" (HON) quality label, and DISCERN scale scores were used to verify their efficiency as quality indicators. Of 874 websites identified, 388 were included. Despite an observed association with higher content quality scores, the HON label fails to predict good content quality websites when used in a multiple regression. Sensibility and specificity of a DISCERN score >40 in the detection of good content quality websites were, respectively, 0.45 and 0.96. The DISCERN is a potential quality indicator with a relatively high specificity. Further developments in this domain are warranted in order to facilitate the identification of high-quality information on the web by patients.

Keywords Internet · Quality indicators · Health care · Consumer · Health-on the net · DISCERN

Introduction

Internet is an important source of information on health [1] and mental health related issues [2]. Reasons to visit medical websites are most frequently to seek information or advice on

Y. Khazaal (✉) · A. Chatton · D. Zullino · R. Khan
Division of Addictology, Geneva University Hospitals, Rue verte, 2, 1205 Geneva, Switzerland
e-mail: yasser.khazaal@hcuge.ch

symptoms, diseases, medical condition or treatments [3, 4]. Prevalence of mental health disorders and addiction as well as the burden of these disorders and possible barriers to treatment [5] may lead people and their relatives to search on-line evidence-based information. This makes it particularly important for websites to present high quality and accurate information which may help people to make informed choices about their treatment.

Unfortunately, it appears from several studies that the quality of information on health and mental health as well as addiction websites is poor [6–8] despite rare exceptions such as bipolar disorder [9].

Content quality indicators are warranted in order to help patients and consumers to judge the content quality of on-line information and to discriminate good sites from others.

Several tools for appraising the quality of health related information for general consumer had been developed. Amongst others, Health on the Net foundation (HON) has issued a code of conduct for medical sites [10, 11] and attributes a quality label taking into account the following points: disclosure of authorship, sources, updating of information, disclosure of editorial and publicity policy, as well as confidentiality.

Several other web-sites quality indicators were defined [12–14] taking into account a variety of factors such as design and aesthetics of the site, readability, dating of information, authority of source, accessibility, disclosure of authors and sponsors. Composite scores based on the sum of several factors [8] or on a overall subjective quality were also proposed [6]. Several recommendations and definitions of good health-related websites were also developed [15, 16].

With few exceptions such as the HON label, most of these possible quality indicators remain complex measures probably more helpful for site owner and designer than for a common consumer.

The DISCERN instrument is a reliable questionnaire developed in order to enable patients and people without content expertise to assess the content quality of information about treatment choice such as the extent to which the information appears unbiased [17, 18]. The instrument comprises 16 items each rated from 1 to 5. Consumer and health professionals DISCERN ratings have been found to be significantly correlated [19]. Despite several studies on its reliability, it remains that data on the validity of DISCERN as indicator of the content quality of health-related websites is still limited.

Furthermore, the DISCERN score predicts content quality in some studies [18–20] but not others [8, 9]. Likewise, the HON label predicts content quality in some studies [20, 21] but not in others [8, 9, 22]. This discrepancy is possibly due to a relatively low rate of sites holding the HON label in several studies. Similarly, DISCERN was rarely applied in conjunction to a tool measuring evidence-based health information to a high number of websites.

Recently, a brief version of the DISCERN was proposed. Sensibility and specificity of a cut-off score was computed for the detection of good content quality websites [23]. To the best of our knowledge, similar data on the DISCERN and the HON are not available to date and could be interesting due to the potential interest of these tools and their frequent use.

Surprisingly, the origin of the sites (university, commercial…) does not predict content quality in several previous studies [8, 9, 22, 24]. So, sites owned by Universities have usually no better content quality than commercial sites. This finding has been considered as possibly due to the small number of sites in several owner categories.

The present study aimed to determine content quality indicators on a large sample of health-related websites and particularly to assess the value of HON label and DISCERN as potential content quality indicators.

Methods

The present study is based on the analysis of data issued from six previous studies [8, 9, 20, 22, 25, 26] which assessed with a standardized tool the general and content quality (evidence-based health information) of health-related websites for the following domains: gambling, alcohol, cocaine, cannabis, bipolar disorder and social phobia. The first aim of these previous studies was to evaluate the content quality, readability, aesthetic, accountability, and interactivity of websites found with key words requests on general search engines. With the exception of bipolar disorder websites, it appears that the content quality of the websites studied was relatively poor (less than average or average). Furthermore these studies assessed aesthetic, readability, interactivity, HON label, DISCERN and global score (the sum of accountability, interactivity, content quality and aesthetic criteria) as potential indicators of content quality. Aesthetic, readability and interactivity were repetitively found as negatively associated with content quality. This results concords with studies from other groups [6, 27] and may reflect various conception of websites (more interactive, less informative vs. more informative, less interactive). The global score appeared as a good content quality indicator. It remains however a complex measure. Results on the HON and DISCERN were variable on these six previous studies. The DISCERN score predicts content quality in some studies [20] but not others [8, 9]. Likewise, the HON label predicts content quality in some studies [20] but not in others [8, 9, 22]. The collection of data assessed similarly for 388 websites offer the opportunity to assess the sensibility and the specificity of HON label and DISCERN as potential useful and simple content quality indicators which was not done in the previous studies with smaller sample sizes.

Typical searches were performed to produce lists of website similar to those generated by a common user with limited internet or medical knowledge.

Selection of Websites

Keyword searches and websites evaluation were done between July 2006 and September 2007 by 3 psychologists and three medical doctors. Keywords related to social phobia, bipolar disorders, pathological gambling as well as cannabis, alcohol and cocaine addiction were entered into popular World Wide Web search engines (Table 1).

Table 1 Keywords and search engines used

Health-related websites	Keywords	Search engines
Social phobia	Social phobia, social anxiety disorder	Google, Yahoo, MSN, AOL, ASK
Bipolar disorder	Bipolar disorder, manic depressive illness	Google, Yahoo
Cannabis addiction	Cannabis addiction, cannabis dependence, cannabis abuse	Google, Yahoo
Cocaine addiction	Cocaine, cocaine addiction and cocaine dependence	Google, Yahoo
Alcohol addiction	Alcohol addiction, alcohol dependence, alcohol abuse	Google, Yahoo, MSN
Pathological gambling	Gambling, pathological gambling, excessive gambling, gambling problem, gambling addiction	Google, Yahoo

The first 20 English language websites forthcoming from each keyword query were examined, as most people rarely search beyond the first 20 retrieved links [7].

Sites were excluded if: inaccessible (invalid address), already reviewed in the current study, containing no information on the topic, requiring access fee, discussion group or open forum, not a site (external links, books or articles) and no information in English.

Evaluation of the Websites

Websites affiliations were divided into five 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) 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 includes most of the health-related websites quality criteria as defined by the Commission of European Communities [15] and the American Medical Association [16].The proforma based on previous studies [6, 12, 20, 28–30], assessed content quality, accountability (Silberg scale) [12], interactivity [30], Abbott's aesthetic criteria [28]and readability[28].

Accountability was assessed using a 9-point scale [6, 12] to rate sites based on criteria of authorship: whether authors and their affiliation and credentials were identified; attribution: whether sources and references were mentioned; disclosure: whether ownership of the site, sponsoring and advertising were disclosed; and currency: whether the date of creation and modification of the site has been specified. For each covered item, 1 point is given to the site.

Interactivity was assessed using an adaptation of Abbot's scale [30] evaluating the presence of audio or video support, a within site search engine, satisfaction and knowledge evaluation questionnaires for users, supporting bodies (forums, discussion rooms), and the possibility to send queries to the webmaster or authors. For each covered item, 1 point is given to the site.

Aesthetic was assessed using Abbott's criteria [30]adapted by Kisely [28] evaluating presence of, diagrams, hyperlinks as well as absence of advertising as well as headings and subheadings. For each covered item, 1 point is given to the site.

Readability was assessed using the Flesch-Kincaid grade level score and the Flesch-Kincaid readability index [28]. The first score evaluates the degree of text reading difficulty with regard to USA school. A score of 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.

Content quality was evaluated on the availability of responses to probable queries. Questions focused on advice for treatment and information concerning the diagnosis, the epidemiology and the complications of each medical condition studied, as it was previously shown that these kind of questions are the most common patients and relatives query on the Internet [3]. The retrieved information was compared to an appropriate guideline for each query: The guidelines of the American Psychiatric Association (2006) for bipolar disorder, cocaine, cannabis and alcohol dependence, the Practice Guidelines for Treating Gambling-related Problems developed by the Massachusetts Council on compulsive gambling, January 2004 and an expert consensus [31] for social phobia.

A content quality score was then defined as the sum of coverage (exhaustibility) and accuracy (correctness) [29]. Similar to the previously described [6, 20], a global score was

defined as the sum of Silberg, interactivity, adapted Abbott's aesthetic criteria and content quality.

Finally, the DISCERN was used. For each study, inter-rater reliability of proforma scores was assessed based on a random sample of sites with at least two trained evaluators.

Raters were trained to use the proforma (with standardized coding process for each component of the proforma) during group sessions before contributing to the evaluation of the websites. Particularly, they were trained to DISCERN according to the guidelines available on the DISCERN website www.discern.org.uk/ which contains an online version of the DISCERN instrument and handbook. Raters completed in group a practice run of a set number of materials which are checked and discussed before contributing to a formal review process. They were similarly trained to the assessment of content-quality according to the specific guidelines chosen. An inter-rater evaluation was then systematically performed and appeared to be good as detailed in the 6 previous studies, particularly for DISCERN (r was always >0.900; $P < 0.01$) and for content quality (r was always >0.800; $P < 0.01$) [8].The inter-rater reliability, as previously shown [22] was also good for the other components of the proforma such as: Silberg (r = 0.841; $P < 0.05$), Flesh Reading ease (r = 0.886; $P < 0.05$), Flesch Kincaid level (r = 0.865; $P < 0.01$), Abbott's esthetic criteria (r = 0.801; $P < 0.05$), and interactivity (r = 0.805; $P < 0.01$).

Analyses

Statistical analyses were performed using SPSS for Windows (version 15.0). Student's *t*-tests were used to compare sites having the HON label and sites without this label. One-way analysis of variance (or its non-parametric counterpart) was also performed to compare sites according to their origin (commercial, university, government, etc.).

Two multiple regression analyses including the HON on one side and the DISCERN on the other side amongst others as predictors were done with a view to predicting the content quality of websites. Receiver Operating Characteristic (ROC) curves were then used to compare the predictive abilities of these two competing models. To this end, the predicted values of each regression were saved and imputed as test variables. The area under the curve (AUC) serves as an indication of the discrimination capability of each regression model: the greater the area, the more the regression model discriminates between sites with poor content quality and sites with good to very good content quality. In other words, the further the curve lies above the reference line, the more accurate the regression model. We prespecified an acceptable AUC ≥ 0.7.

Depending on the medical subject studied in the previous studies, content quality maximum scores are 20, 24 or 28. A standardized 28 points content quality score was obtained as follows: two weighted factors of 2/7 and 1/7, respectively, are multiplied by this maximum and the result added to the current value of the content quality. For instance, the standardized value for a site with current content quality score of 16 and initial maximum score of 20 is 24 (16 + 2/7*28) whereas this value becomes 20 (16 + 1/7*28) if the initial maximum score is 24. Based on previous studies [23], a cut-off point of 20/28 has been chosen taking into account the similar cut-off used in school or university (20/28 = 14.3/20 = 71.4/100). This cut-off is higher than 14, the mean score for a minimal and mostly right information. It has been chosen to delineate sites with poor content quality and sites with good to very good content quality, it being understood that the last one scored 20 and above.

Before proceeding with the prediction, a factor analysis was carried out in order to reduce a large number of correlated predictors, namely the Silberg scale, the Abbott's

aesthetic criteria, the Flesch-Kincaid grade level, the Flesch-Kincaid readability index and the Interactivity scores with the aim of incorporating only a small subset of uncorrelated variables.

To test whether Factor analysis is satisfactory, we used the Kaiser–Meyer–Olkin (KMO) and Bartlett's tests to measure the sampling adequacy. The KMO measure should be equal or greater than 0.5 to proceed with Factor analysis, while the Barlett's test of sphericity should be statistically significant at the 0.05 level, meaning that the correlation matrix is not an identity matrix.

Two criteria for deciding how many factors to retain were observed: the Kaiser rule and the Cattell scree plot obtained by principal components extraction.

The Kaiser rule states that all factors with Eigen values greater than 1 should be retained. Although this rule seems conservative or old fashioned, one has to bear in mind that Eigen values less than 1 originate from factors that account for less variance than the original variables.

The Cattell scree plot, i.e. a graph which shows the magnitude of the Eigen values versus the components serve to confirm these factors: the rule being that factors above the inflexion point of the slope must be retained. 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.

The factor analysis being completed, the extracted significant factors, the origin of the sites, the HON alternately with the DISCERN were analyzed through two multiple linear regressions for the prediction of content quality. These two regressions are presented as model 1 and model 2, respectively. Dummy-coding scheme was adopted for categorical variables. The assumptions of independent normally-distributed errors and constant variance of errors were checked by studying the residuals from the model. Potential collinearity among the independent variables was studied through the variance inflation factor (VIF) value and goodness of fit of the model was assessed through the coefficient of determination: the adjusted R^2 statistic.

For all analyses, a significance level of $P \leq 0.05$ was used.

Results

We reviewed 874 websites. After application of the exclusion criteria, the data of 388 web sites on alcohol dependence (n = 103), bipolar disorder (n = 34), cannabis addiction (n = 57), cocaine addiction (n = 61), gambling (75) and social phobia (n = 58) were analysed together. Means and standard deviation scores are reported in Table 2.

Twenty-one per cent (21%) of the sites had the HON label. Results of the comparison between the sites holding the HON label with those without the label are reported in Table 3. HON label websites have higher scores on content quality, accountability (Silberg score), interactivity and modified global scores.

Results of the comparison of sites categories (university, commercial…) are reported in Table 4. No differences were found on content quality score nor DISCERN. Nevertheless, significant differences appear on accountability, aesthetic and readability criteria.

Regarding the variable reduction for the prediction, the KMO measure and the Barlett's test are all satisfactory. Two factors that have been extracted by principal components with a varimax rotation account for almost 66% of the variance. The Flesch-Kincaid grade level and the Flesch-Kincaid readability index load on factor 1, thereafter named "readability".

Table 2 Websites' general and content quality indicators, mean ± SD, (N = 388)

Frequency of site origin	
Government	35
Organisation	93
University	59
Commercial	171
Individual	13
Unknown	17
Presence of the HON (health on the net) logo %	
No	79
Yes	21
Silberg scores (0–9), mean (SD)	4 (2.3)
Interactivity scores (0–6)	1.7 (1.2)
Abbott aesthetic criteria scores(0–4)	2.7 (0.8)
Flesch reading ease scores (0–100)	44.8 (13.6)
Flesch-Kincaid education scores (1–12)	9.8 (2.8)
Standardized content quality scores (0–28)	16.3 (5.9)
Modified Global scores (0–17)[a]	8.5 (3)
DISCERN scores (16–80)	35.2 (11.2)

[a] The sum of aesthetic criteria, interactivity and Silberg scores (with the exclusion of content quality scores)

Table 3 Comparison of instrument scores by HON label

	With HON label (n = 81)	Without HON label (n = 307)	t	P-value
Silberg scores	5.2 (1.8)	3.7 (2.3)	−6.45	<0.0005
Interactivity scores	2.2 (1.2)	1.6 (1.2)	−3.69	<0.0005
Aesthetic criteria scores	2.6 (0.7)	2.8 (0.9)	2	0.05
Flesch reading ease scores	41.2 (14.5)	45.8 (13.2)	2.54	0.02
Flesch-Kincaid education scores	10.3 (2.8)	9.7 (2.9)	−1.5	n.s.
Content quality scores	17.73 (6.2)	15.9 (5.7)	−2.36	0.02
Modified Global scores	10 (2.5)	8 (3)	−5.76	<0.0005
DISCERN scores	37.3 (11)	34.6 (11.2)	−1.9	n.s.

The Silberg and Interactivity scores load on factor 2, called "accountability and interactivity".

The model with HON label, origin of the sites (government site as the reference category), readability, accountability-interactivity as predictors and content quality as the outcome using enter method (model 1) yields the following observations. The overall F statistic of the regression was significant ($F_{(7,376)} = 2.1$ and $P = 0.04$), rejecting the null hypothesis that the population multiple correlation is 0 and the adjusted R^2 statistic showed that only 2% of the total variation in the outcome variable was explained by this model. There was no statistically significant predictor as shown in Table 5.

A new prediction model with the DISCERN variable shows that the overall F statistic was highly significant ($F_{(7,376)} = 44.2$ and $P < 0.0005$). The adjusted R^2 statistic revealed that 44% of the total variation in the outcome variable is explained by this model. Unlike the previous model, two predictors (factor 2 and DISCERN) significantly explained the content quality. The results are presented in Table 6.

Table 4 Comparison of instrument scores by origin of site

	Government	Organisation	University	Commercial	Individual	Unknown	F or χ^2	P
Silberg	5.1 (2.5)	4.6 (2.2)	3.9 (2)	3.7 (2.2)	3.2 (2.4)	3 (2.4)	4.49	0.001
Interactivity	1.9 (0.9)	1.9 (1.1)	1.5 (1.1)	1.7 (1.3)	1.3 (0.9)	1.6 (1.5)	9.3a	n.s.
Abbott Aesthetic	3.3 (0.6)	2.9 (0.7)	2.3 (0.9)	2.7 (0.8)	2.2 (0.7)	2.8 (0.7)	43.2a	<0.0005
Flesch reading	39.9 (13)	41 (13.7)	49.9 (13)	46 (12.8)	51.7 (16.4)	41 (12)	5.5	<0.0005
Flesch-Kincaid	10.6 (2.7)	10.5 (2.5)	8.4 (2.8)	9.8 (2.8)	8 (3.9)	11.4 (2.4)	7	<0.0005
Content quality	17.2 (6.1)	16.6 (5.2)	15.2 (6.8)	16.4 (5.6)	14.5 (7.8)	16.7 (6.4)	1.9a	n.s.
Global scores	27.4 (7.7)	26 (6.3)	22.8 (7.7)	24.6 (6.4)	21 (9.5)	24 (7)	10.8a	n.s.
Modified Global scores	10.3 (2.9)	9.4 (2.8)	7.6 (2.9)	8.2 (2.9)	6.6 (3)	7.4 (3.4)	7.3	<0.0005
DISCERN	39 (12.1)	36.5 (10.2)	34.7 (12.4)	33.7 (10.9)	39.2 (14.4)	33.5 (8.7)	9.96a	n.s.

[a] χ^2 value obtained by non parametric Kruskal–Wallis test

Table 5 Model 1: prediction of the content quality of web sites dealing with information on health: summary of a multiple linear regression analysis with the HON variable, amongst others, as predictor

| Predicting variable	Unstandardized coefficients		Standardized coefficients	t	Sig.	95% Confidence interval for B	
	B	SE	β			Lower bound	Upper bound
Constant	16.23	1.03		15.82	0.000	14.21	18.24
Factor 1	0.53	0.31	0.09	1.72	0.087	−0.08	1.13
Factor 2	0.50	0.32	0.08	1.55	0.122	−0.13	1.12
Individual site	−0.60	1.48	−0.03	−0.40	0.688	−3.52	2.32
University site	−0.10	1.11	−0.01	−0.09	0.926	−2.28	2.07
Organisation site	−0.8	1.16	−0.01	−0.07	0.944	−2.37	2.21
Other sites	−0.97	1.30	−0.06	0.75	0.455	−3.53	1.58
HON label	1.47	0.75	0.10	1.94	0.053	−0.02	2.95

Table 6 Model 2: prediction of the content quality of web sites dealing with information on health: summary of a multiple linear regression analysis with the DISCERN variable, amongst others, as predictor

Predicting variable	Unstandardized coefficients		Standardized coefficients	t	Sig.	95% Confidence interval for B	
	B	SE	β			Lower bound	Upper bound
Constant	3.20	1.09		2.93	0.004	1.05	5.35
Factor 1	−0.00	0.23	0.00	−0.01	0.993	−0.46	0.46
Factor 2	−0.76	0.25	−0.13	−3.08	0.002	−1.25	−0.28
Individual site	−1.08	1.12	−0.05	−0.96	0.336	−3.28	1.13
University site	0.73	0.84	0.06	0.87	0.386	−0.92	2.37
Organisation site	0.17	0.87	0.01	0.19	0.848	−1.55	1.89
Other sites	−1.12	0.98	−0.07	−1.15	0.253	−3.05	0.80
DISCERN	0.37	0.02	0.70	17.04	0.000	0.33	0.41

Once the two regressions are finalized, the respective predicted values are saved and subsequent ROC analyses may be conducted. We examined the performance and efficiency of each model. Model 1 did not meet the pre-specified AUC threshold of 0.7. Its exact value was 0.63 (CI ∈ [0.57, 0.69]). Considering that good or very good sites have a content quality score of 20 and above, its sensibility and specificity were 0 and 100%, respectively, at this specified cut-off. Regarding model 2, the AUC was 0.83 (CI ∈ [0.79, 0.88]) and its sensibility and specificity 45 and 96%, respectively, using the same cut-off. Hence, every 4 sites out of 100 on average would be incorrectly identified as good to very good. As shown with ROC curves (Fig. 1) and Table 7, DISCERN is a better discriminator than the HON label.

Discussion

The present study aimed to determine content quality indicators of mental health related internet sites. In the present study, content quality was not associated with origin of sites
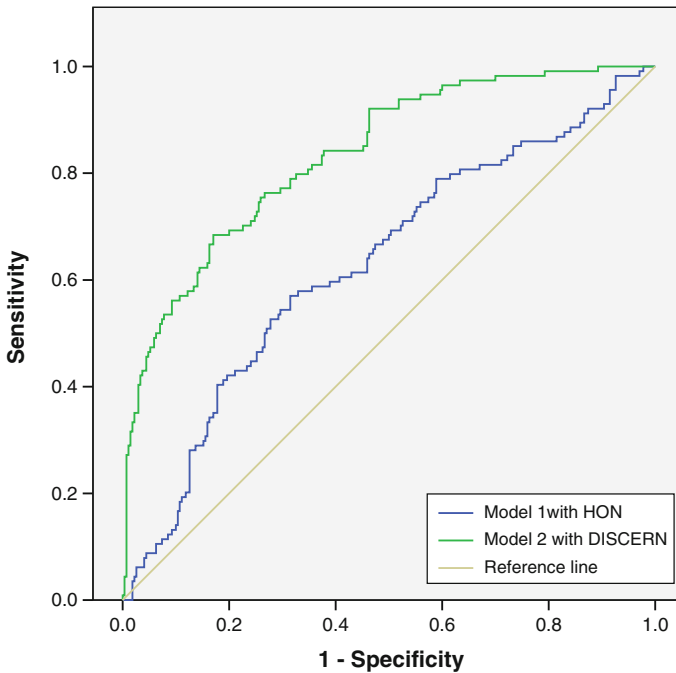
Fig. 1 ROC curve

Table 7 ROC: performance of 2 content quality predictive models at cut point = 20		Model 1: with HON	Model 2: with DISCERN
	AUC	0.63	0.83
	Sensitivity	0	0.45
	Specificity	1	0.96
[a] Cannot be calculated since the model with HON identified no good sites	Positive predictive value	–[a]	0.77
	Negative predictive value	0.70	0.81

neither with the HON label. Content quality was however positively associated with DISCERN. Although sites holding the HON label appear to have higher content quality score than sites without the HON label, the predictive ability of this label to detect good content quality websites was not demonstrated in this study. In a multiple regression analysis including the DISCERN, it appears that DISCERN is positively associated with content quality, whereas factor 2 (accountability and interactivity) was negatively associated with content quality. This last result is in accordance with previous studies showing a negative relation between factors such as readability, accountability and interactivity and content quality [18, 27]. It was previously hypothesized 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 [9].

Surprisingly, origin of sites including University websites was not associated with content quality. One plausible explanation might be the lack of financial means available

for this purpose. Considering the complexity of the creation of a health-related web site, this is not a surprise.

In the present study, the HON label does not appear as a good content quality indicator despite the higher content quality scores obtained by the sites holding the HON label. This is possibly due to the fact that the HON label is more in relation to an ethical chart than to content quality in itself. The DISCERN score was associated with content quality which is in accordance with several previous studies [18, 20].

A DISCERN score >40 out of 80 seems to be a potentially good indicator of good content quality websites with a relatively high specificity (0.96) and an average sensibility (0.45). Using this cutoff, only 4% of all poor content quality websites were incorrectly identified as good to very good ones. This results is quite similar, a little bit better than the ones obtained with the 6-items version of the DISCERN, the cut-off value of the Brief DISCERN (sensibility 0.357 and specificity 0.945) [23].

The present study has several limitations. It only gives account of the situation on bipolar disorder, social phobia, cannabis, cocaine, alcohol and gambling related websites related websites between July 2006 and September 2007. So, the conclusion may be not true of the whole health-related websites domains. General search engines were used rather than medical search engines which, are however more commonly used by general population, and lead to quite similar results [32]. The study may not reproduce behaviors of all Internet consumers who possibly use variable search methods in regard to keywords and search engines used. A further limitation is that same coders have assessed websites with DISCERN, HON and content quality. It however remains that the inter-coder reliability was high on a random sample of websites especially for DISCERN and for content quality considered separately (evaluation of DISCERN without any measure of content quality and conversely). It is furthermore difficult to have DISCERN coders fully blind to content quality, because raters have to read text before encoding.

Furthermore, in the present study, websites were evaluated by professionals. This aspect may limit the extension of the results to laypersons. It was however found in a previous study that consumer and health professionals DISCERN ratings are significantly correlated [19]. Further studies may however help to connect content quality and patient's treatment through informed choices by having consumers themselves assessing the information and indicate treatment-choices, a process which may be also mediated by other factors than the evaluation of content-quality.

Nonetheless, this study brings to evidence on a large sample of health-related one interesting content quality indicator.

Conclusion

In conclusion, the DISCERN is a potential quality indicator with a relatively high specificity.

References

1. Powell J, Clarke A: The WWW of the World Wide Web: Who, what, and why? Journal of Medical Internet Research 4(1):e4, 2002
2. Powell J, Clarke A: Internet information-seeking in mental health: Population survey. British Journal of Psychiatry 189:273–277, 2006

3. Shuyler KS, Knight KM: 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(4):e24, 2003

4. Khazaal Y, Chatton A, Cochand S, et al.: Internet use by patients with psychiatric disorders in search for general and medical informations. Psychiatric Quarterly 79:301–309, 2008

5. Chartier-Otis M, Perreault M, Belanger C: Determinants of barriers to treatment for anxiety disorders. Psychiatric Quarterly 81(2):127–138, 2010

6. Griffiths KM, Christensen H: Quality of web based information on treatment of depression: Cross sectional survey. BMJ 321(7275):1511–1515, 2000

7. Eysenbach G, Powell J, Kuss O, et al.: Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 287(20):2691–2700, 2002

8. Khazaal Y, Chatton A, Cochand S, et al.: Quality of web-based information on cocaine addiction. Patient Education and Counseling 72(2):336–341, 2008

9. Morel V, Chatton A, Cochand S, et al.: Quality of web-based information on bipolar disorder. Journal of Affective Disorders 110(3):265–269, 2008

10. Boyer C, Selby M, Appel RD: The health on the net code of conduct for medical and health web sites. MedInfo 9 Pt 2:1163–1166, 1998

11. Boyer C, Selby M, Scherrer JR, et al.: The health on the net code of conduct for medical and health Websites. Computer in Biology and Medicine 28(5):603–610, 1998

12. Silberg WM, Lundberg GD, Musacchio RA: Assessing, controlling, and assuring the quality of medical information on the internet: Caveant lector et viewor–Let the reader and viewer beware. JAMA 277(15):1244–1245, 1997

13. Jadad AR, Gagliardi A: Rating health information on the Internet: navigating to knowledge or to Babel? JAMA 279(8):611–614, 1998

14. Gagliardi A, Jadad AR: Examination of instruments used to rate quality of health information on the internet: Chronicle of a voyage with an unclear destination. BMJ 324(7337):569–573, 2002

15. eEurope 2002: Quality criteria for health related websites. Journal of Medical Internet Research 4(3): E15, 2002

16. Winker MA, Flanagin A, Chi-Lum B, et al.: Guidelines for medical and health information sites on the internet: Principles governing AMA web sites. American Medical Association. JAMA 283(12): 1600–1606, 2000

17. Charnock D, Shepperd S, Needham G, et al.: DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. Journal of Epidemiology and Community Health 53(2):105–511, 1999

18. Griffiths KM, Christensen H: The quality and accessibility of Australian depression sites on the World Wide Web. Medical Journal of Australia 176(Suppl):S97–S104, 2002

19. Griffiths KM, Christensen H: Website quality indicators for consumers. Journal of Medical Internet Research 7(5):e55, 2005

20. Khazaal Y, Fernandez S, Cochand S, et al.: Quality of web-based information on social phobia: A cross-sectional study. Depression and Anxiety 25(5):461–465, 2008

21. Fallis D, Fricke M: Indicators of accuracy of consumer health information on the Internet: A study of indicators relating to information for managing fever in children in the home. Journal of the American Medical Informatics Association 9(1):73–79, 2002

22. Khazaal Y, Chatton A, Cochand S, et al.: Quality of web-based information on cannabis addiction. Journal of Drug Education 38(2):97–107, 2008

23. Khazaal Y, Chatton A, Cochand S, et al.: Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. Patient Education and Counseling 77(1):33–37, 2009

24. Khazaal Y, Chatton A, Cochand S, et al.: Quality of web-based information on alcohol dependence. Drugs: Education, Prevention, and Policy 17(3):248–260, 2010

25. Khazaal Y, Chatton A, Cochand S, et al.: Quality of web-based information on pathological gambling. Journal of Gambling Studies 24(3):357–366, 2008

26. Khazaal Y, Chatton, A., Cochand, S., Coquard, O., Fernandez, S.,Khan, R., Zullino, D.: Quality of web based information on alcohol dependence. Drugs: Education, Prevention & Policy 17(3):248–260, 2009

27. Hargrave D, Bartels U, Lau L, et al.: Quality of childhood brain tumour information on the internet in French language. Bull Cancer 90(7):650–655, 2003

28. Kisely S, Ong G, Takyar A: 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(1):85–91, 2003

29. Nilsson-Ihrfelt E, Fjallskog ML, Blomqvist C, et al.: Breast cancer on the Internet: The quality of Swedish breast cancer websites. Breast 13(5):376–382, 2004

30. Abbott VP: Web page quality: Can we measure it and what do we find? A report of exploratory findings. Journal of Public Health Medicine 22(2):191–197, 2000
31. Van Ameringen M, Allgulander C, Bandelow B, et al.: WCA recommendations for the long-term treatment of social phobia. CNS Spectrum 8(8 Suppl 1):40–52, 2003
32. Ilic D, Bessell TL, Silagy CA, et al.: Specialized medical search-engines are no better than general search-engines in sourcing consumer information about androgen deficiency. Human Reproduction 18(3):557–561, 2003

Author Biographies

Yasser Khazaal, MD is a specialist in psychiatry and psychotherapy in the Division of Addictology at the University Hospitals of Geneva, Switzerland. He has more than 90 scientific publications, mainly on cognitive behavior therapy and pharmacotherapy as well as internet-related behaviors.

Anne Chatton, MA is statistician. She has a number of publications in the field of mental health.

Daniele Zullino, MD is a specialist in psychiatry and psychotherapy. He is the head of the Division of Addictology at the University Hospitals of Geneva, Switzerland. He has more than 200 scientific publications.

Riaz Khan, MD is a specialist in psychiatry and psychotherapy in the Division of Addictology at the University Hospitals of Geneva, Switzerland.