

Documental Quality of Websites Concerning Physical Activity, Lifestyle and Sedentarism Available on the Internet

Marcos Bagrichevsky^{1*} , Paulo Roberto Vasconcellos-Silva²

¹Department of Medicine, Regional University of Blumenau, Blumenau, Brazil

²Federal University of State of Rio de Janeiro, Rio de Janeiro, Brazil

Email: *marcos_bagrichevsky@yahoo.com.br

How to cite this paper: Bagrichevsky, M. and Vasconcellos-Silva, P.R. (2019) Documental Quality of Websites Concerning Physical Activity, Lifestyle and Sedentarism Available on the Internet. *Health*, 11, 1684-1692. <https://doi.org/10.4236/health.2019.1112127>

Received: December 1, 2019

Accepted: December 28, 2019

Published: December 31, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Objectives: To analyze the documental quality of 389 websites in Portuguese about physical activity, healthy lifestyles and sedentary lifestyles found on the Brazilian version of the general search engine *Google*. **Methods:** The documental quality of the 389 websites was estimated based upon the following parameters: 1) a combination of quality criteria from the Health Information Locator (LIS—OPS/BIREME) and those from Chile’s Pontifical Catholic University, organized into 17 variables; 2) uniformity of reference criteria (Vancouver); 3) association between the presence of authorship and a higher number of the quality criteria being fulfilled. We also studied the ranking of the results presented by *Google* in addition to attributes connected to the websites’ target audience, the types of content, their sponsors and country of origin. **Results:** Of the 389 websites studied, 111 links were not active (28.53% CI 95% [24.05 - 33.02]) and none of the websites in the sample met all of the 17 quality variables. Authored websites displayed remarkable differences in quality when compared to those which did not identify their authors. **Conclusions:** Faced with the issue of the proliferation of websites with questionable quality content, and the fact that the ranking of results interferes directly in the internal evaluation of content relevance, we propose that public-health research institutions cooperate with web-searching developers to improve the website-positioning formula, in which the “identified authorship” criterion should play a major role in the ranking system.

Keywords

Health Information, Websites, Documental Quality, Physical Activity

1. Introduction

Advances in information technology have brought about undeniable gains in the

field of therapy and disease diagnosis. On the other hand, while “efficacy” discourses underlie the production of scientific content, there is also a profusion of content directed at the mass consumption of health-related products [1]. Various studies have related that the Internet fills a huge gap as the most popular place of reference about these issues [2] [3]. For the producers of the information sites, the “Web” does not pose restrictions on the suitability of any person or body who wants to add content about health promotion, whether the motivation is preventive or commercial. This adds to the difficulties in controlling the quality of health information put on the Internet [4] and makes assessing this issue of the Internet highly interesting [5] [6].

The rapid growth in websites dedicated to the health sciences causes difficulties when it comes to ensuring and judging the quality of the information they offer. There are many proposals by Health Organizations that seek criteria to identify high-quality information on the Internet. Because of the disparity in technical knowledge between experts and the general public, accurate information on these subjects is not always clear and easily understandable. It becomes increasingly difficult to discern the information that is updated and produced with technical rigor from the inaccurate, or obsolete information that may, eventually, mislead the public [5].

Nowadays it is not uncommon for patients to turn to their doctors with a great volume of information obtained on the Internet regarding their diseases. In about 15% of cases, the diagnosis and treatment offered by the health professional are questioned based upon information from dubious sources [7]. It is becoming increasingly frequent for patients to resort to the Web in order to find a “virtual diagnosis” or “elixir” to alleviate symptoms of illness. As there are no widely accepted certification filters, the user cannot evaluate the quality of health information content and tends to make the mistake of thinking that, because it is on the Internet, all the information there has unquestionable scientific value.

The questionable content of the specialized information on websites is also growing in the field of health promotion. The rhetoric ranges from acclaiming physical activity as a simple measure for solving complex health problems to the legitimization of products on the beauty market. The paradox is that the imperative to lead a “healthy lifestyle” seems more related to physical appearance and moral values than to health itself [8] [9].

Whether it focuses on self-treatment or on health promotion, the ease of finding a website on the Internet (given that this is connected to its greater visibility) is directly related to its credibility. This fact boosts the belief that “friendly (understood as familiar) is also dependable”. The website positioning algorithms (page rank) and *Google’s* “sample fallacies”, described below, directly interfere with the positioning of the results that connect to the websites, deciding the sequence in which they appear, which directly influences the user’s judgment of their pertinence and relevance.

This study sought to analyze (through criteria postulated by widely accepted institutions in this field) the documental quality of websites whose content deals

with combating a sedentary lifestyle and promoting healthy lifestyles (through physical activity), found on *Google*. We also checked the sites' main attributes, their visibility, as well as the positive association between the presence of authorship and compliance with quality indicators.

2. Methods

Three hundred and eighty-nine (389) Portuguese language websites found on the Brazilian version of the general search engine *Google* (<http://www.google.com.br/>) were evaluated (Search date: Nov 15, 2012), through “advanced search” and the option “include all words”, using as metadata (search terms) the phrases: “physical activity”, “lifestyle” and “sedentarism”.

It is well-known that the *Google* search engine tends to change the results displayed from one search to the next or on different computers [10] [11]. In order to avoid mistakes caused by changes in results and retain the “positioning” (original order of the results), the results were recorded in pdf format (option: “100 results per page”) with a hyperlink. Exclusion criterion: pages that requested payment before consultation or that limited access to registered users.

An Excel 2003[®] spreadsheet was used for storing the data and hyperlinks, mathematical calculations and creating tables and figures. The statistical analysis procedures were performed using the program SPSS[®] version 15.0 for Windows.

In the documental quality evaluation we used a check-list of 17 quality indicators for websites (shown in **Table 1**), developed from the combination of criteria from two sources: 1) information quality criteria indicators from the Health Information Locator—LIS (OPS/BIREME), Dublin Core (LIS—Health Information Locator <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>); and 2) the health information quality criteria from Chile's Pontifical Catholic University, *Pontificia Universidad Católica* [12] [13] [14] [15]. We also considered as documental quality criteria the six reference uniformity indicators (Vancouver) established by the International Committee of Medical Journal Editors (ICMJE, <http://www.icmje.org/>): presence of an author; title; editor or responsible institution; date of creation; date of update; and country.

The following features were also analyzed: inactive links; target audience; and objectives (academic, commercial, governmental, etc.).

We also studied the correlation (Pearson method) between “authorship presence” and “greater compliance with these information quality criteria”.

3. Results

3.1. Description of the Websites

All the 389 results displayed by *Google* were studied, and 111 inactive links were found (28.5% CI 95% [24.0 - 33.0]). From the other 278 results, 129 examples linked to academic articles about physical activity and sedentary lifestyle (46.4% CI 95% [40.5 - 52.3]); 78 to commercial content (28.0% CI 95% [22.8 - 33.3]);

Table 1. Check-list—indicators related to documental quality.

-
- 1) **Authorship**—people responsible for the displayed contents.
 - 2) **Promoting body**—institution that supports the website.
 - 3) **Endorsement**—relevant professional or academic institutions that give guarantees and endorse the site.
 - 4) **Coherence of the title and the content.**
 - 5) **Dates of creation and Web publication.**
 - 6) **Date of update**—last change.
 - 7) **Links**—functioning of the three first links, if they exist (from left to right and top to bottom on the homepage).
 - 8) **Coherence of links**—relationship and pertinence of the external *sites* to the subject of the text.
 - 9) **Existence of contact details**—means for contacting those responsible: addresses, telephone numbers, electronic mail, etc.
 - 10) **Contact validity**—response from the person responsible for the *website* when receiving a personal request for information.
 - 11) **Help**—visible presence of help that makes understanding the website easier.
 - 12) **Information management**—option of saving or easily generating a file in pdf format for printing.
 - 13) **Navigability**—use of the homepage and related pages without the need to use specific software.
 - 14) **Usability**—absence of difficulties for people with special needs (visual, hearing, physical or cognitive impairment.) Preferably with the W3C certification.
 - 15) **Certification**—legitimacy of structural quality or of contents confirmed by certifying bodies: HONcode, etc.
 - 16) **Conflicts of interest**—declaration of any secondary influence, usually economic or personal.
 - 17) **Objectivity**—clarity of ideological, moral, ethical, religious, commercial and personal opinions.
-

30 to governmental or administrative information (10.8% CI 95% [7.1 - 14.4]); and 41 to other subjects (14.7% CI 95% [10.6 - 18.9]). One hundred and forty-eight (148) homepages directed at the general public were identified (53.2% CI 95% [47.4 - 59.1]); 80 allowed the general public restricted access (28.8% CI 95% [23.4 - 34.1]); and 50 (17.99% CI 95% [13.5 - 22.5]) offered access options for the general public and professionals. Brazil was the country of origin of 265 websites (95.3% CI 95% [92.8 - 97.8]). For the remaining 13 (4.7% CI 95% [2.2 - 7.2]) it was not possible to establish their origin.

3.2. Documental Quality

No website fulfilled all the 17 indicator criteria in **Table 1**, and none of them indicated the date of update or declaration of conflicts of interest (**Table 2**). A maximum score of 14 indicators fulfilled (1 case, or 0.3% CI 95% [0.0 - 1.1]) and a minimum of 0 (3 cases 1.1% CI 95% [0.0 - 2.3]) were observed. A mean of 9.2 (± 0.1) and a median of 10 variables were met. In relation to the fulfillment of the six Vancouver uniformity criteria, no site met all these variables—the mean was 3.5 ± 0.1 and the median was 4, with a maximum of 5 (29 cases 10.4% CI 95% [6.8 - 14.0]) and minimum of 0 (4 cases 1.4% CI 95% [0.0 - 2.8]).

3.3. Authorship Presence

Among the pages studied, the presence of authorship may be positively correlated to six quality indicators (**Table 3**): presence of a sponsoring body; title

Table 2. Omitted items in the pages about physical activity, lifestyle and sedentarism.

Attribute	Percentage	IC 95%
Author not identified	27.7	22.44 - 32.96
Organization body not identified	30.2	24.82 - 35.61
No institutional validation	32.3	26.87 - 37.87
Lack of title coherence	1.8	0.24 - 3.36
No creation date	78	73.19 - 82.92
No update date	100	100 - 100
Invalid links (first three)	9.3	5.93 - 12.78
Lack of coherence between the links	80.3	75.42 - 85.20
Non-existence of contact details	14.7	10.58 - 18.92
Invalid contact details	29.5	23.73 - 35.34
Absence of help texts	12.9	9.00 - 16.90
Lack of information management	71.	66.28 - 76.88
Difficult to navigate	1.08	0.00 - 2.29
Compromised usability	2.9	0.91 - 4.84
Lack of certification	99.7	98.94 - 100
Absence of conflict of interests declaration	100	100 - 100
Lack of objectivity	76.9	72.03 - 81.93

Table 3. Correlation between authorship and further documental quality criteria in web-sites about physical activity, lifestyle and sedentarism.

	<i>p</i>
Responsible body*	<0.001
Endorsement (Institutional)	0.759
Title coherence*	<0.001
Creation date	0.721
Update date	Not calculated: no differences
Link validity (first three)	0.580
Link coherence	0.767
Existence of contact details*	0.033
Validity of contact details*	0.047
Help	0.108
Information management*	0.028
Navigability*	0.005
Usability	0.153
Certification	0.535
Conflict of interests declaration	Not calculated: no differences
Objectivity	0.931

*Significant level at 0.05.

coherence; existence and validity of contact details; information management and navigability. Likewise, significant differences between the averages of the variable quality criteria sum ($p < 0.001$) were highlighted.

4. Discussion

From the point of view of documental quality, it is possible to state that the websites analyzed displayed unsatisfactory results. Additionally, the number of inactive links on homepages was higher than described in other works [16] [17] [18], which may indicate the presence of a significant number of unstable and, perhaps, untrustworthy sources. The documentation uniformity criterion may contribute to the evaluation of this picture, due to the need to identify the origin of the sources and when they were updated: failures in websites references make it difficult to verify their trustworthiness [19] [20] [21] [22].

In order to find out whether the information disseminated by a periodical and a general or technical book is up-to-date, it is essential to know the date of its creation or publication as well as, in some cases, the date on which it was last updated. Equally essential to the credibility of texts, especially in specialized technical literature, is the possibility of contact with the person responsible for the published document, which may not be specified in about a third of the sample analyzed in our investigation. It is also worrying to note that among the homepages studied, more than three quarters did not record the date when they were first published and none recorded the dates they were updated. A technical book would be useless without this information.

Considering such results and in the face of the high number of websites dedicated to the health sciences, the question arises: how can we ensure and judge the quality of so much information that deals with such complex subjects and need specialized and up-to-date technical validation? There are similar works in the literature that use indicators comparable to the ones in this study [23]. Nevertheless, all acknowledged that users needed a high education level—that is, it is necessary to have previous education for users to ensure they can benefit from the resources available on the Internet [24] [25] [26] [27].

In order to analyze tools that categorize the reliability and validity of health information, Gagliardi and Jadad [28] carried out a criteria review study for the validation of specialized websites on this subject. Despite a large sample, the results indicated that only five of them had instructions for use. Additionally, none of them offered information about the reliability and validity of the measurements. Of the 98 tools identified at the beginning of the investigation, 47 had disappeared at the end. In summary, the review concluded that inadequate tools were being used for evaluating health information on the Web.

The poor quality of information found on the Internet is a problem that is also connected to changes in indicators of quality, which are well known for printed literature. However, there are not enough analogous indicators for choosing a good text in the unstable world of new media [24]-[29]. In the present work,

authored websites offered some degree of accountability in the sense that their authors may be held partially responsible for their contents. On the other hand, health professionals, who are in a position to better advise non-specialists what is the best way in which to select relevant and up-to-date material, still do not have efficient tools for indicating trustworthy information sources to the general public. In the current picture, the fact that health information consumers are teaching (and diagnosing [30]) themselves presents potential dangers and, in this context, we would highlight the role of the institutions with healthcare remits who are able to promote studies that allow us to validate the information published on the Web.

5. Conclusion

We concluded that the documental quality of information found in Brazilian websites is very poor. However, it should be admitted that even fulfillment of all the variables presented here, would not by itself totally guarantee the quality of the contents of the analyzed websites. Authored websites offer some degree of accountability in the sense that their authors may be held partially responsible for their contents. Faced with the issue of the proliferation of websites with questionable quality content, and the fact that the ranking of results interferes directly in the internal evaluation of content relevance, we propose that public-health research institutions cooperate with web-searching developers to improve the website-positioning formula, in which the “identified authorship” criterion should play a major role in the ranking system. This concept should be the object of further analysis by specialists in scientific documentation, who would be responsible for recommending selection criteria, above all the existence of website authorship, for the ranking of results. This would facilitate the internal review and evaluation (those that the users themselves use in order to select their texts) in terms of pertinence and relevance.

Collaborators' Participation

The authors developed and organized all parts of the text in conjunction.

Acknowledgements

To the National Council for Scientific Research and Development—CNPq for funding the research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Beck, U. (2002) *La sociedad del riesgo global*. Siglo Veintiuno, Madrid.
- [2] Sillence, E., Briggs, P., Harris, P. and Fishwick, L. (2007) *Going Online for Health*

- Advice: Changes in Usage and Trust Practices over the Last Five Years. *Interactive Computing*, **19**, 397-406. <https://doi.org/10.1016/j.intcom.2006.10.002>
- [3] Kerr, C., Murray, E., Stevenson, F., Gore, C. and Nazareth, I. (2006) Internet Interventions for Long-Term Conditions: Patient and Caregiver Quality Criteria. *Journal of Medicine and Internet Research*, **8**, e13. <https://doi.org/10.2196/jmir.8.3.e13>
- [4] Pallen, M. (1995) Guide to the Internet: The World Wide Web. *British Medical Journal*, **311**, 1552-1556. <https://doi.org/10.1136/bmj.311.7019.1552>
- [5] Eysenbach, G., Powell, J., Kuss, Q. and 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. <https://doi.org/10.1001/jama.287.20.2691>
- [6] Vasconcellos-Silva, P.R., Castiel, L.D., Griep, R.H. and Zanchetta, M. (2008) Cancer Prevention Campaigns and Internet Access: Promoting Health or Disease? *Journal of Epidemiology Community Health*, **62**, 876-881. <https://doi.org/10.1136/jech.2007.069054>
- [7] Murray, E. (2003) The Impact of Health Information on the Internet on Health Care and the Physician-Patient Relationship: National U.S. Survey among 1500 US Physicians. *Journal of Medicine and Internet Research*, **5**, e17. <https://doi.org/10.2196/jmir.5.3.e17>
- [8] Bagrichevsky, M. and Santos, D.S. (2018) "Shameless" Sedentarism: Individual Responsibility for Health? *Psychology*, **9**, 760-772. <https://doi.org/10.4236/psych.2018.94048>
- [9] Lupton, D. and Chapman, S. (1995) A Healthy Lifestyle Might Be the Death of You: Discourses on Diet, Cholesterol Control and Heart Disease in the Press and among the Lay People. *Sociology of Health and Illness*, **17**, 477-494. <https://doi.org/10.1111/1467-9566.ep10932547>
- [10] Culebras-Fernández, J.M., Sanz-Valero, J. and Wanden-Berghe, C. (2008) The Quality of Nutrition Websites: A Necessity! *Proceedings of the 8th Congresso Regional de Informação em Ciências da Saúde y 5th Reunião de Coordenação da Biblioteca Virtual de Saúde*, Rio de Janeiro, 16-19 September 2008, e1-e5.
- [11] Sanz-Valero, J., Wanden-Berghe, C., Nolasco, A., Bernebeu-Mestre, J. and Guadiola, W.B.R. (2008) Information about Nutrition in Google: Sequels of Page-Rank. *Proceedings of the 30th Congress of the European Society for Clinical Nutrition and Metabolism*, Florence, 13-16 September 2008, 4 p.
- [12] LIS-Regional [Website]. LIS: Modelo da Biblioteca Virtual em Saúde. BIREME/OPS, São Paulo. <http://lis.bvs.br/lis-Regional/>
- [13] Biblioteca Nacional de Ciencias de la Salud (BNCS) [Sede Web] (2005) Metodología LIS—Localizador de informação en Salud: Criterios para selección de fuentes de informação en saúde disponibles en Internet. BNCS, São Paulo. http://files.sld.cu/liscuba/files/2010/01/criterios-de-seleccion_liscuba.pdf
- [14] National Library Board [Website] (2008) Dublin Core Metadata Initiative, DCMI Metadata Terms. National Library Board, Singapur. <http://dublincore.org/documents/dcmi-terms>
- [15] Pontificia Universidad Católica de Chile (PUC) [sede Web]. Criterios generales para evaluar recursos en Internet. PUC, Santiago de Chile. http://www.puc.cl/sw_educ/gnosis/H/gnosish4.htm
- [16] Dellavalle, R.P., Hester, E.J., Heilig, L.F., Drake, A.L., Kuntzman, J.W., Graber, M., et al. (2003) Information Science. Going, Going, Gone: Lost Internet References. *Science*, **302**, 787-788. <https://doi.org/10.1126/science.1088234>

- [17] Belda Albero, A. (2008) La qualidate estructural y documental de los websites sobre nutrición en internet. MSc. Tesis, Centro de Investigación en Economía y Gestión da Salud, Universidad Politécnica, Valencia.
- [18] Martín Martínez, B. (2007) Valoración da qualidate de las páginas web en gastroenterología, hepatología y nutrición infantil. PhD Tesis, Universidad de Barcelona, Facultad de Medicina, Barcelona.
- [19] Johnson, K.R., Hester, E.J., Schilling, L.M. and Dellavalle, R.P. (2004) Addressing Internet Reference Loss. *The Lancet*, **363**, 660-661. [https://doi.org/10.1016/S0140-6736\(04\)15608-0](https://doi.org/10.1016/S0140-6736(04)15608-0)
- [20] Schilling, L.M., Kelly, D.P., Drake, A.L., Heilig, L.F., Hester, E.J. and Dellavalle, R.P. (2004) Digital Information Archiving Policies in High-Impact Medical and Scientific Periodicals. *JAMA*, **292**, 2724-2726. <https://doi.org/10.1001/jama.292.22.2724>
- [21] Sanz-Valero, J., Wanden-Berghe, C. and Castiel, L.D. (2006) Retorno a Vancouver. *Gaceta Sanitaria*, **20**, 251-252. <https://doi.org/10.1157/13088859>
- [22] Sanz-Valero, J., Wanden-Berghe, C., García de Lorenzo, A., Gil, A., Pérez da Cruz, A., Sánchez-Álvarez, C., *et al.* (2007) Las normas de publicación de Nutrición Hospitalaria; un futuro para la revista en beneficio de sus autores. *Nutrición Hospitalaria*, **22**, 1-3.
- [23] Fallis, D. and Fricke, M. (2002) 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**, 73-79. <https://doi.org/10.1136/jamia.2002.0090073>
- [24] Ferguson, T. (2002) From Patients to End Users. *British Medical Journal*, **324**, 555-556. <https://doi.org/10.1136/bmj.324.7337.555>
- [25] Ritchie, L., Tornari, C., Patel, P.M. and Lakhani, R. (2016) Glue Ear: How Good Is the Information on the World Wide Web? *The Journal of Laryngology & Otology*, **130**, 157-161. <https://doi.org/10.1017/S0022215115003230>
- [26] Ferrari, C.C., Spampatti, S., Leotta, A., Rausei, S., Rovera, F., Boni, L., Inversini, D., Carcano, G., Dionigi, G. and Dionigi, R. (2013) Web-Based Information on Intraoperative Neuromonitoring in Thyroid Surgery. *International Journal of Surgery*, **11**, S40-S41. [https://doi.org/10.1016/S1743-9191\(13\)60012-6](https://doi.org/10.1016/S1743-9191(13)60012-6)
- [27] Priyanka, P., Hadi, Y.B. and Reynolds, G.J. (2018) Analysis of the Patient Information Quality and Readability on Esophagogastroduodenoscopy (EGD) on the Internet. *Canadian Journal of Gastroenterology and Hepatology*, **2018**, Article ID: 2849390. <https://doi.org/10.1155/2018/2849390>
- [28] Gagliardi, A. and Jadad, A.R. (2002) Examination of Instruments Used to Rate Quality of Health Information on the Internet: Chronicle of a Voyage with an Unclear Destination. *British Medical Journal*, **324**, 569-573. <https://doi.org/10.1136/bmj.324.7337.569>
- [29] Hirsch, M., Aggarwal, S., Barker, C., Davis, C.J. and Duffy, J.M.N. (2017) Googling Endometriosis: A Systematic Review of Information Available on the Internet. *American Journal of Obstetrics and Gynecology*, **216**, 451-458. <https://doi.org/10.1016/j.ajog.2016.11.1007>
- [30] Bartolomé, A. (2004) Aprendizaje potenciado por la tecnología: Razones y diseño pedagógico. In: Martínez Sánchez, F. and Prendes Espinosa, M.P., Eds., *Nuevas Tecnologías y Educación*, Pearson Educación, Madrid, 215-234.