# Quality of online information on breast cancer treatment options

Article (Accepted Version)

**Quality of online information on breast cancer treatment options.**

Nadia Arif and Pietro Ghezzi

*Brighton & Sussex Medical School, Falmer, BN19RY, UK*

Correspondence: P. Ghezzi, Brighton & Sussex Medical School, Falmer, BN19RY, UK. E-mail: p.ghezzi@bsms.ac.uk; Phone: +44(0)1273873112

## Highlights

- Analysis of 200 websites returned searching Google for “breast cancer treatment options”.
- Most websites give a fair representation of surgical and medical treatments.

  Complementary and alternative medicines described more frequently by commercial websites.
- Ranking by Google prioritizes non-profit organizations and government websites.

**ABSTRACT**

Offering breast cancer patients treatment choice has become a priority as the involvement of patients in the decision-making process is associated with improved physical and psychological outcomes. As the Internet is increasingly being used by patients as a source of medical information, it is important to evaluate the quality of information relating to breast cancer on the Internet. We analysed 200 websites returned by google.co.uk searching “breast cancer treatment options” in terms of their typology and treatment options described. These were related to standard measures of health information quality such as the JAMA score and the presence of quality certifications, as well as readability.

We found that health portals were of higher quality whilst commercial and professional websites were of poorer quality in terms of JAMA criteria. Overall, readability was higher than previously reported for other conditions, and Google ranked websites with better readability higher. Most websites discussed surgical and medical treatments. Few websites, with a large proportion being of commercial typology, discussed complementary and alternative medicine. Google ranked professional websites low whilst websites from non-profit organizations were promoted in the ranking.

**Key words**. Internet; Google; health information quality; information quality; online; breast cancer.

**Abbreviations**. CAM, Complementary and Alternative Medicine; ECQC, European Commission quality criteria; FK, Flesh-Kincaid; HON, Health on the Net Foundation; HIQ, Health Information Quality; IQ, Information Quality; IRR, Interrater Reliability; JAMA, Journal of American Medical Association; SERP, Search Engine Results Page; SMOG, Simple Measure of Gobbledygook; URL, Uniform Resource Locator.

**Funding sources**. This research was funded by the Brighton & Sussex Medical School as part of an independent research project for 4th year medical students.

## INTRODUCTION

The Internet is an important source of medical information for patients; 35% of the US population [1], and over 50% in the EU [2], searched for health information online. Earlier studies were concerned that patients could find low-quality information [3], and thus several assessment tools were developed to evaluate health information quality (HIQ), including the Journal of American Medical Association (JAMA) criteria [3] and the Health on the Net Foundation seal (HONcode) [4]. Ease of readability is another parameter evaluated in addition to trustworthiness [5-7].

Breast cancer is the commonest cancer among women. Treatment options include surgical, medical and complementary and alternative medicine (CAM) [8]. Mastectomy and breast conservation surgery with radiotherapy are the most common management options [9], and offering patients treatment choice has become a priority [10-12].

Sixty-three percent of cancer patients use the Internet for information, with a higher rate of use (73%) in breast cancer patients [13], mostly using search engines, primarily Google [14, 15]. Cancer patients use the Internet to verify information received from their doctors and to develop questions to discuss with them, as well as to seek alternative treatments [13]. A 2014 study on breast cancer patients found that “improvement of knowledge obtained through personal research on the Internet, books and other media” is an independent predictor of an active role in the choice of therapy [16]. Early studies have warned that breast cancer patients may be basing their decisions on inaccurate or incomplete information [17-19]. As summarised in Table 1, several studies have analysed the HIQ of websites on breast cancer using different methods.

**Table 1. Literature on IQ of breast cancer and the assessment tools used**

| Search query | No. of websites | HIQ tool | Readability | Content analysis | Ref. |
|---|---|---|---|---|---|
| Breast cancer symptoms, breast cancer care, breast cancer stage, breast cancer survival, breast cancer signs | 289 English | JAMA | - | | [20] |
| Breast cancer | 29 Swedish | ECQC | - | Coverage, correctness | [19] |
| Breast cancer, childhood asthma, depression, obesity | 18 English and 7 Spanish | - | Yes | Coverage, correctness | [17] |
| Breast cancer | 184 English | JAMA, HONcode | - | Coverage | [18] |
| Cancer, breast cancer, breast cancer information | 10 English | ECQC | - | Coverage, correctness | [21] |
| Breast cancer surgery, breast cancer treatment, mastectomy, lumpectomy | 45 English | DISCERN | | | [22] |
| Breast reconstruction post mastectomy | 71 English | HONcode, University of Michigan Consumer Health Website Evaluation Checklist | Yes | | [23] |

A study measuring the completeness of online information on breast cancer found that for some important topics the relevant clinical information had been mentioned only briefly [17]. A more recent study found that although government, charity and formal educational websites had very high accuracy, inaccurate information on breast cancer was prevalent on the Internet [20].

The aim of this study was to assess websites on breast cancer treatment options and to ascertain the visibility given by Google to websites discussing CAM. This is particularly important to investigate as online health information can have significant implications on the patient's decision-making regarding treatment options. The search query "breast cancer treatment options" is also very sensitive to news reports, as shown by a spike in 2013 following Angelina Jolie's mastectomy announcement [24].

Google was used as it is the primary search engine for over 80% of users [25]. The intrinsic dimensions of HIQ were assessed using the JAMA criteria, HONcode and ease of readability, in addition to basic content analysis on the specific type of treatment mentioned, whether medical, surgical or CAM. Because patients rarely browse beyond the first 10 websites returned by a Google search engine result page (SERP) [26], we also analysed how websites were ranked by Google.

## METHODS

### Data collection

A search on 'breast cancer treatment options' was conducted in September 2016 on Google.co.uk. Search history, cookies and caches were cleared to avoid the possible influence of prior browsing history. The first 200 URLs of the SERP were transferred onto a spreadsheet and visited. Sample size is based on our previous studies indicating that it is powered enough to detect differences in the composition of the SERP [27-30]. Inaccessible websites (requiring registration or subscriptions), duplicates, and those containing no information were then excluded. Figure 1 summarises how the websites were analysed.

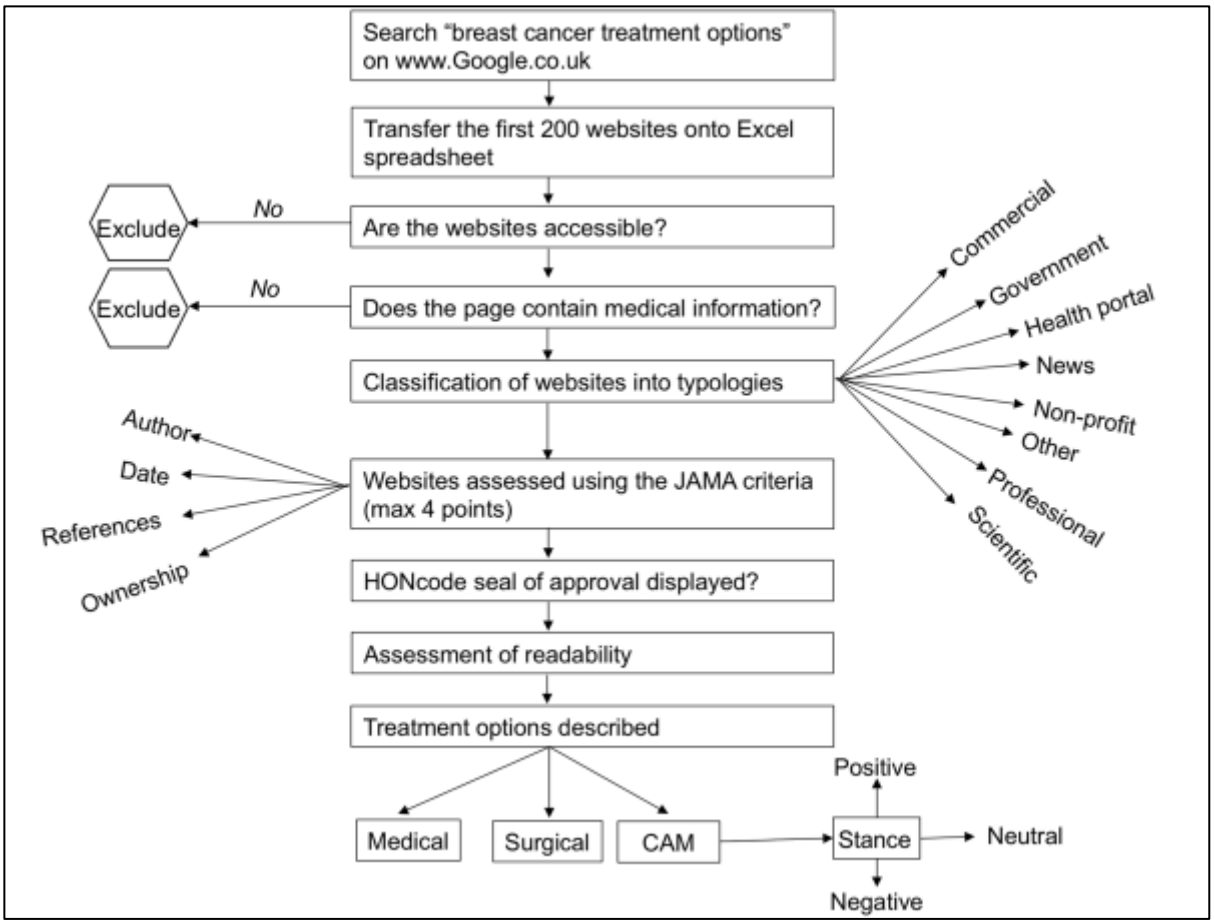


**Figure 1. Data collection and processing.**

### Analysis of websites

Websites were analysed according to the criteria below. In assessing websites, if a criterion was not visible on the initial webpage, the 3-click rule was used, where if a specific feature could not be found within three clicks, the website was given a score of 0 for that criterion [27].

1. Typology. Two investigators categorised all the websites into distinct typologies as described in Table 2 [27, 28].

**Table 2. Examples of website typologies**

| Typology | Description | Example |
|---|---|---|
| Commercial (C) | Websites that buy, sell or provides a service for a fee with the aim of making a profit. | www.roche.com<br>www.healthcare.siemens.com |
| Government (G) | Website created, managed or regulated by an official governmental body. | www.nhs.uk<br>www.canceraustralia.gov.au |
| Health portal (HP) | Website with a search function that contains health information on a variety of health topics. | www.webmd.com<br>www. patient.info |
| News (N) | Website from newspapers, magazines or TV created for the distribution of news and information. | www.time.com<br>www.telegraph.co.uk |
| Non-profit (NP) | Organisation with charitable/ supportive/ educational services that are not established for the purpose of profit-making. | www.cancerresearchuk.org<br>www.macmillan.org.uk |
| Others (O) | Websites that do not fit into any of the other typology classifications. | www.messageboard.4hcm.org/forum<br>www.ibcsupport.org/treatment.html |
| Professional (P) | Websites created by health professionals, experts and professional organisations. | www.mayoclinic.com<br>www.health.clevelandclinic.org |
| Scientific journals (S) | Scientific journals online or academic publishing | www.sciencedirect.com<br>www.oncology.jamanetwork.com |

Interrater reliability (IRR) between the two investigators' classification was then calculated. There were 181 agreements (96%) between the two investigators, which was deemed 'very good' (Cohen's kappa coefficient, 0.95). The agreement varied between 86% for commercial websites and 100% for government and scientific websites. Where there was a disagreement in the classification, the websites were revisited and a consensus was achieved through discussion.

2. JAMA score. The websites were evaluated for the following four features: authorship, attribution, disclosure and indication of date. A score of 1 was assigned for the presence of each of these criteria, therefore websites were scored from 0 to 4. JAMA scores were assigned independently by the two investigators and the scores compared to calculate the IRR. Of the 188 websites assessed, there were

only seven disagreements (96% agreement). The strength of this IRR was also considered to be ‘very good’ (Cohen’s kappa coefficient, 0.95). Disagreements were resolved by the investigators through a discussion and reaching a consensus.

3. HONcode certification. Websites were searched to determine whether a HONcode certification was displayed.

4. Readability. An online readability test tool was used [31]. The reading grade levels of all the websites were calculated using two different readability formulas, the Flesh-Kincaid (FK) and the Simple Measure of Gobbledygook (SMOG). While the FK grade considers the average sentence length and the average number of syllables per word [32], the SMOG formula takes also into account the number of polysyllabic words in 30 sentences [32]. A lower grade indicates a readability suitable for lower age groups, and therefore easier to read. Eight websites could not be investigated as they were not accessible to the readability software.

5. Treatment options. We noted the treatment options discussed (medical, surgical or CAM), and whether clinical trials were mentioned. Although 21 websites mentioned CAM, five were not counted as CAM because they maintained a negative stance on it.

Statistical analysis was performed using Graphpad Prism 7.0 (GraphPad Software, San Diego, USA); the statistical tests used are described in the text.

## RESULTS

### Composition of the SERP and ranking by Google

Of the 188 URLs in the search, the most frequent typologies were professional (42%) and non-profit (17%) (Table 3).

**Table 3. Distribution of websites by typology**

| Typology | Total websites (n=188) | Top 10 (n=10) |
|---|---|---|
| **Commercial** | 14 (7%) | - |
| **Government** | 12 (6%) | 2 (20%) |
| **Health portal** | 11 (6%) | - |
| **News** | 14 (7%) | - |
| **Non-profit** | 32 (17%) | 7 (70%)** |
| **Others** | 11 (6%) | - |
| **Professional** | 78 (42%) | 1 (10%)* |
| **Scientific** | 16 (9%) | - |

Number of websites in each typology in the entire SERP and in the top 10 websites returned by Google. *P<0.05, ** P<0.0005 significantly different from the frequency in the total SERP by Fisher's test.

In the top 10 results, Google gives greater visibility to non-profit and government websites. There are also significantly more non-profit websites in the total top 10 (70%) compared to the rest of the SERP (17%). Conversely, professional websites are significantly underrepresented in the top 10 websites returned.

**Analysis of JAMA score and HONcode certification**

Fig.2a shows the median JAMA score of the total websites assessed was 2 (IQR: 0, 4), with health portals, news and scientific journals having the highest average JAMA score and professional and commercial websites the lowest. There was no significant difference in the JAMA score between the top 10 and the remaining websites (Fig.2b).

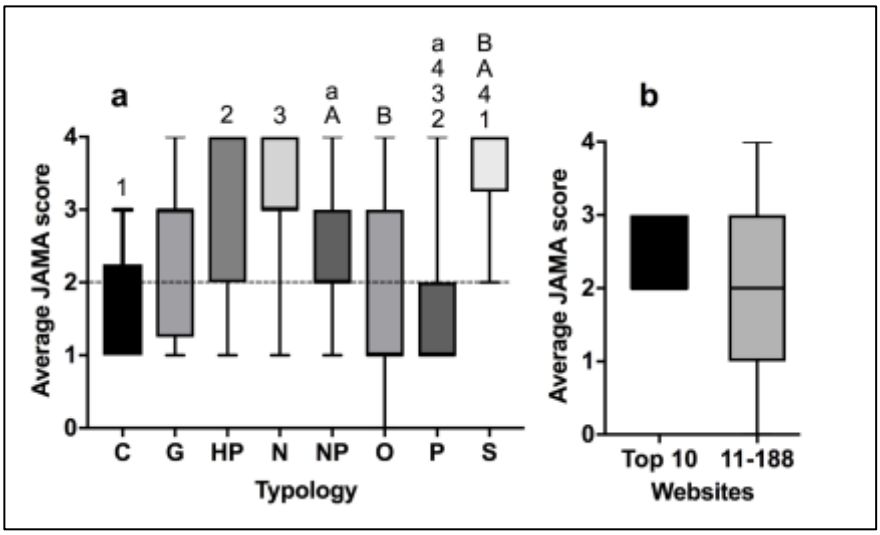


**Figure 2. JAMA score of websites from different typologies (a) and Google ranking (b)**. Numbers indicate the median with the interquartile range. (a) JAMA score by typology in the whole SERP. Number of websites in each typology are as in Table 3. Values bearing the same symbols are significantly different from each other using the Kruskal-Wallis multiple comparisons test corrected for multiplicity using statistical hypothesis testing (numbers, P <0.0001; lower case letters, P <0.001; capital letter, P<0.05). Dotted line represents the median JAMA score of all websites. b) Average JAMA score for websites 1-10 and 11-188.

As a JAMA score ≥ 3 is considered high quality [18], we also analysed the number of websites in each typology meeting this criterion. The results shown in Fig.3 confirm the pattern observed in Fig.2.

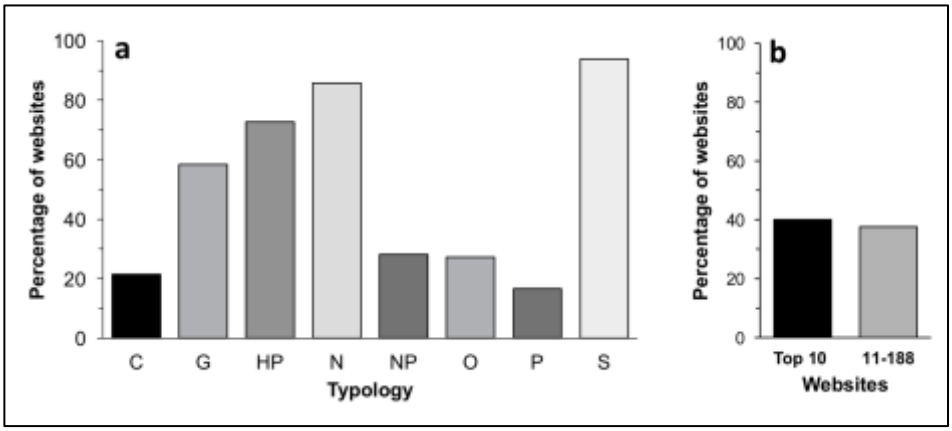


**Figure 3. Percentage of websites in each typology with a JAMA score ≥ 3 by website typology (a) or Google ranking (b).** Number of websites: C, 3/14; G, 7/12; HP, 8/11; N, 12/14; NP, 9/32; O, 3/11; P, 13/78; S, 15/16.

Only 13 out of 188 (7%) websites displayed a HONcode, health portals accounting for eight of them. In fact, 73% of the health portals displayed the HONcode. None of the websites in the top 10 returned by Google displayed a HONcode certification.

The JAMA score correlated with the presence of the HONcode seal: JAMA score was median 3 (IQR:1.5, 4.0) in the 13 websites with the HONcode and median 2 (IQR:1.0, 3.0) in the remaining 175 websites (P=0.0406 using two-tailed Mann-Whitney test).

**Analysis of readability**

Fig. 4 shows the readability of websites as assessed using the FK (panels a, b) and SMOG (panels c, d) grading. The mean FK grade for the total websites was 8.5 (95% CI 7.9-9.1) and the mean SMOG for the 180 websites was 7 (95% CI 6.8-7.2). We could not find any significant difference among the different typologies, except for "other" websites scoring a higher SMOG grade than other groups. However, the SMOG grade of the top 10 websites was significantly lower (better readability) than that of websites 11-188 (Fig.4d).

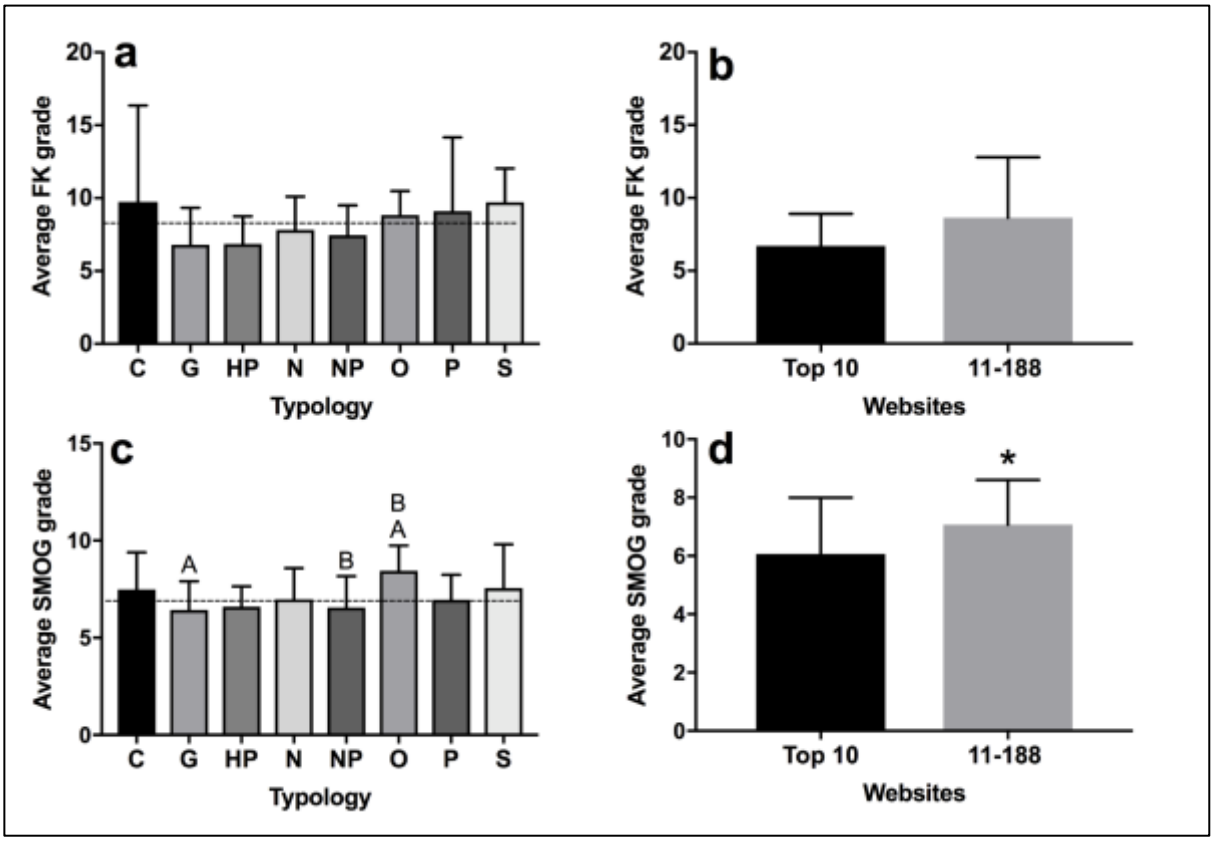


**Figure 4. Website readability.** (a) Average FK grades of the different typologies; (b) average FK grades of the top 10 (n=10) and the remaining websites (n=170); (c)

average SMOG grades of the different typologies; (d) average SMOG grades of the top 10 websites (n=10) and the remaining websites (n=170). Values bearing the same letter are significantly different from each other (P<0.05 by Tukey's test with correction for multiple comparison using statistical hypothesis testing). *P<0.05 vs top 10 websites by Student's *t*-test.

**Correlation of readability with information quality score and ranking by Google**

We also analysed whether the readability of websites correlates with the JAMA score or the presence of the HONcode seal. As expected, the FK and SMOG grades correlate strongly (P<0.0001, r=0.6902, using a two-tailed Spearman test, n=180).

There was no correlation between the JAMA score and either the FK grade (P=0.7385) or the SMOG grade (P=0.7415). However, HONcode certified websites (n=13) had a lower FK grade (better readability) than the 167 websites without HONcode certification (HONcode+, median 6.3 (IQR: 6.0, 7.3); HONcode-, median 7 (IQR: 6.1, 8.0); P=0.0094 using two-tailed Mann-Whitney test). This difference was not observed with the SMOG grade (median was 6.0 in HONcode+ and 6.1 in HONcode-, P=0.19).

We also analysed the association of readability with Google ranking. The FK grade was significantly lower (better readability) in the top 10 websites, median 5.9 (IQR: 5.3, 8.4) compared to the remaining websites, median 7.0 (IQR: 7.0, 9.6); P=0.0253 using two-tailed Mann-Whitney test. This difference did not reach the significance for the SMOG grade (top 10 websites had a median of 5.8 (IQR: 4.4, 7.0) while the remaining websites had a median of 7.0 (IQR: 6.1, 8.0); P=0.0575).

**Treatment options discussed by websites**

We evaluated the different treatment options for breast cancer that were discussed. Fig.5 shows that 155 (82%) websites mentioned medical, 124 (66%) websites mentioned surgical, and 21 (11%) websites mentioned CAM treatment options, with several websites mentioning more than one option. Most websites mentioned both medical and surgical treatment options (n=107) with only 13 mentioning all three categories.

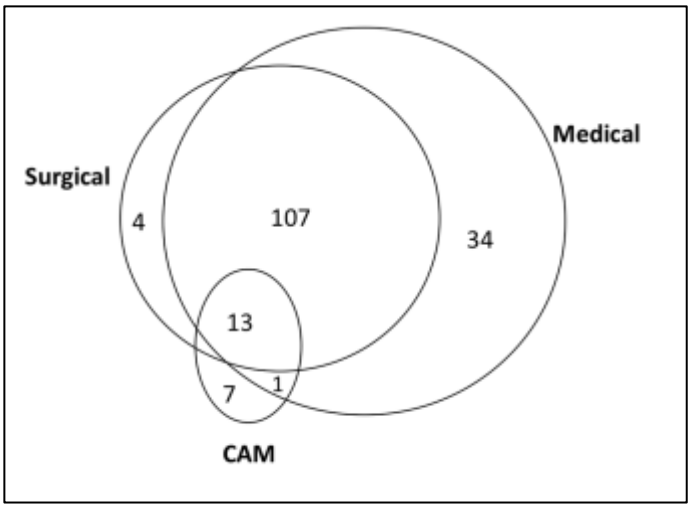


**Figure 5. Websites discussing the different breast cancer treatment options.**

We next analysed the distribution of the typologies within the different treatment option categories recommended by websites against the expected distribution of those typologies. The expected typology composition is the distribution of the typologies in the total 188 websites (Table 3). As shown in Table 4, a significantly higher proportion than expected of commercial websites discussed CAM (25%).

**Table 4. Composition of observed typology in the three different treatment option categories**

| | Observed | | | Expected |
|---|---|---|---|---|
| Typology | **Medical** | **Surgical** | **CAM** | **Composition of typology in total websites** |
| **Commercial** | 10 (6.5%) | 7 (6%) | 4 (25%) ***** | 14 (7%) |
| **Government** | 10 (6.5%) | 8 (6%) | 1 (6.25%) | 12 (6%) |
| **Health** | 10 (6.5%) | 8 (6%) | 1 (6.25%) | 11 (6%) |
| **News** | 11 (7%) | 4 (3%) ****** | 1 (6.25%) | 14 (7%) |
| **Non-profit** | 29 (19%) | 28 (23%) | 5 (31.25%) | 32 (17%) |
| **Others** | 7 (4.5%) | 6 (5%) | 1 (6.25%) | 11 (6%) |
| **Professional** | 65 (42%) | 58 (47%) | 3 (18.75%) | 78 (42%) |
| **Scientific** | 13 (8%) | 5 (4%) ****** | 0 (0%) | 16 (9%) |

Numbers indicate the number of websites discussing each treatment option category. Numbers in parenthesis show the percentage of the treatment options category that the typology contributes to. Five websites that discussed CAM were excluded because they maintained a negative stance and did not recommend it as a treatment option. Values bearing *, $P<0.05$; **, $P<0.01$ are significantly different by Fisher test compared to expected percentage of that typology in the total search.

From the results in Table 4, it also appears that significantly fewer news articles and scientific journals than expected discuss surgery as a treatment option.

Only 19 websites (10%) discussed clinical trials, and half of these were professional (n=10), four non-profit, three government, one scientific journal and one “other” websites. Finally, we found no significant difference on comparing the websites for the treatment options mentioned and their readability (not shown).

## DISCUSSION

In agreement with other studies [27], we found that health portals and scientific journals consistently score better than other typologies using standard HIQ criteria such as JAMA and HONcode certification.

Although the JAMA score and HONcode do not measure the accuracy of the information, a previous study analysing breast cancer website content found that educational websites, encompassing scientific journals and health portals, were more accurate [20, 33]. This is likely due to review boards and policies for publication, that are in place for websites with these affiliations. The fact that these categories have high JAMA scores suggests that the JAMA criteria is a good proxy indicator for content accuracy. We found that health portals were also the easiest to read, along with government and non-profit websites. Not surprisingly, the technical content of scientific journals resulted in the lowest readability.

Breast cancer information on the Internet, as assessed in this study, appears to be more readable compared to other health information on the Internet. Studies have found that the readability of information on Parkinson’s disease was on average FK grade 12 and material on lateral epicondylitis grade 12 [5, 6], while a study evaluating the readability of patient education material from surgical subspecialties found information to be at a high reading level, between 9 and 17 grade levels [7]. Other studies on cancer have found a similar trend [34]. Although breast cancer websites have good readability, they may still not be understood by the average patient as, to be understood by 75% of the population, readability should be at a sixth-grade level [35].

Furthermore, we found a better readability, in terms of FK grades, for HONcode-certified websites and for the top 10 websites in the Google SERP, although this was not statistically significant for the SMOG score. Although we found the two measures to correlate very well, in agreement with another study [36], they are different in how they are derived as has been previously described. The SMOG grade is often considered better for the purpose of health information [37, 38] but was suggested to be less accurate for grades <6 [38].

Content analysis indicated that most websites discuss medical and surgical treatments, with 107/188 mentioning both, thus providing a good coverage of the therapeutic options. Subgroup analysis of the treatments mentioned show that commercial websites were three times more likely to discuss CAM, in agreement with studies on HIQ of websites on antioxidants [30] and influenza prevention [28], thus confirming the trend that commercial websites are more likely to describe therapies outside evidence-based medicine.

In terms of newsworthiness, medical treatments received proportionally more attention than surgery. It would be interesting to see how many of the articles are journalistic and how many are simply echoing press releases of pharmaceutical companies or research institution promoting their work, as have been noted elsewhere [39, 40].

Only 10% of the websites discussed clinical trials. It is well recognised that clinical trials are crucial for improving patient outcomes with cancer and methods of patient accrual are a focus of debate in the oncology community [13]. A study has shown that 23.5% of patients with cancer have used the Internet to find information on clinical trials [13], therefore the Internet can potentially be a significant asset in encouraging patients to enrol in these trials.

Another element of analysis was the Google ranking of websites in the SERP, which is important in determining which information will reach patients. We found that Google gives higher visibility to non-profit and government websites. Contrary to the widespread biased view that Google promotes commercial websites, we could not find any of the 12 commercial websites in the top 10 hits returned by Google, in agreement with our previous reports analysing search results on influenza prevention [28], antioxidants [30], and migraine therapy [29], where we found that commercial websites are ranked low by Google. The reason for this is unclear as the algorithm used by

Google to rank websites is not published. However, it is unlikely that this involves a content analysis and probably reflects a different structural organization of websites of different typologies. For instance, it is possible that lacking transparency indicators (author, date, references to sources) is important as two typologies not showing in the top 10 results (professional and commercial) show a below-average JAMA score, and the JAMA score seems to be, on average, higher in the top 10 websites (Fig.4).

From performing a sub-analysis of the JAMA score components in the present study, we found that “currency” was a criterion met by all the top 10 websites but only observed in 43% of the remaining 178 websites. However, one should be careful in assuming that currency contributes to the Google ranking because we did not observe this in other studies on health information with the same sample size (not shown).

Overall, around 40% of websites had a JAMA score considered good (≥3). A previous study, analysing 45 breast cancer websites using the DISCERN score (which includes references and currency) reported that 31% mentioned the sources and 53% the date [22]. In comparison, the present study found that 35% of the 188 websites mentioned the source of information and 47% the date; authorship was present in 36% and ownership of the website by 97%. While this shows a remarkable consistency in the information quality on breast cancer obtained with different search terms and search engines.

Readability is clearly another IQ criterion that is easily assessed by a machine, and we also noticed that readability is better (i.e. a lower grade level) in the top 10 websites. Overall, the mean FK and SMOG grades in the websites analysed in this study (8.5 and 7.0, respectively) are lower than that reported in a 2013 study of websites returned by google.com [41]. It is possible that readability of websites improved in the last four years but we also need to bear in mind that we have used the local UK version of Google, rather than google.com. Furthermore, we only analysed the webpage returned by Google rather than several articles of the same website.

Further limitations include the use of only one search query. This could be mitigated by the fact that we analysed a large sample of websites but using different search terms might give different results, particularly in terms of Google ranking.

The other limitation is that we used a local search engine (google.co.uk) and did not address websites in other languages. There are many differences between countries

not only for the language but also for the Internet usage, as the percentage of EU population that used the Internet daily 2016 varied between 42% and 92%, with an average of 71% [42].

Another major limitation of the study is that we have investigated a sample of websites and their visibility in terms of Google ranking. However, although the top 10 websites have a higher visibility, they may not be equally read by information seekers, and only questionnaires or studies using eye-tracking devices or web tracking would identify which websites are actually read.

We conclude that the quality of information relating to breast cancer on the Internet is variable, with health portals having higher quality and commercial and professional websites being of poorer quality in terms of standard criteria. However, the fact that professional websites had lower JAMA score confirms that it is not a predictor of the scientific quality of the content. On the other hand, it is reassuring that the main search engine does not rank commercial or low-quality websites highly.

Although the vast majority of websites inform patients on medical and surgical treatments, with few describing only CAM, patients may still stumble upon non-scientifically oriented websites discouraging them from following recommended therapies. Therefore, and given the high levels of Internet use amongst breast cancer patients and its implications, we recommend that healthcare professionals take greater responsibility in evaluating various websites in terms of scientific accuracy. This would allow them to signpost and guide their patients towards high quality health information online. Finally, guidance is also needed to disseminate information on clinical trial outreach strategies in order to influence enrolment of more patients on clinical trials.

## REFERENCES


[1] Fox S, Duggan M. Health online 2013. Health. 2013.

[2] Eurostat. Internet access and use statistics - households and individuals. Eurostat Statistics Explained2017.

[3] 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. 1997;277:1244-5.

[4] Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health Websites. Comput Biol Med. 1998;28:603-10.

[5] Dy CJ, Taylor SA, Patel RM, McCarthy MM, Roberts TR, Daluiski A. Does the quality, accuracy, and readability of information about lateral epicondylitis on the internet vary with the search term used? Hand (N Y). 2012;7:420-5.

[6] Fitzsimmons PR, Michael BD, Hulley JL, Scott GO. A readability assessment of online Parkinson's disease information. J R Coll Physicians Edinb. 2010;40:292-6.

[7] Hansberry DR, Agarwal N, Shah R, Schmitt PJ, Baredes S, Setzen M, et al. Analysis of the readability of patient education materials from surgical subspecialties. Laryngoscope. 2014;124:405-12.

[8] Choices N. Breast cancer (female) - Treatment 2016.

[9] Caldon LJ, Walters SJ, Ratcliffe J, Reed MW. What influences clinicians' operative preferences for women with breast cancer? An application of the discrete choice experiment. Eur J Cancer. 2007;43:1662-9.

[10] Caldon LJ, Walters SJ, Reed MW. Changing trends in the decision-making preferences of women with early breast cancer. Br J Surg. 2008;95:312-8.

[11] Stewart MA. Effective physician-patient communication and health outcomes: a review. CMAJ. 1995;152:1423-33.

[12] Fallowfield L, Hall A, Maguire P, Baum M, A'Hern R. A question of choice: results of a prospective 3-year follow-up study of women with breast cancer. The Breast. 1994;3:202-8.

[13] Castleton K, Fong T, Wang-Gillam A, Waqar MA, Jeffe DB, Kehlenbrink L, et al. A survey of Internet utilization among patients with cancer. Support Care Cancer. 2011;19:1183-90.

[14] McLeod J, Yu I, Ingledew PA. Peering into the Deep: Characterizing the Internet Search Patterns of Patients with Gynecologic Cancers. J Cancer Educ. 2017;32:85-90.

[15] Nguyen SK, Ingledew PA. Tangled in the breast cancer web: an evaluation of the usage of web-based information resources by breast cancer patients. J Cancer Educ. 2013;28:662-8.

[16] Taioli E, Joseph GR, Robertson L, Eckstein S, Ragin C. Knowledge and prevention practices before breast cancer diagnosis in a cross-sectional study among survivors: impact on patients' involvement in the decision making process. J Cancer Educ. 2014;29:44-9.

[17] Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. JAMA. 2001;285:2612-21.

[18] Meric F, Bernstam EV, Mirza NQ, Hunt KK, Ames FC, Ross MI, et al. Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites. Bmj. 2002;324:577-81.

[19] Nilsson-Ihrfelt E, Fjällskog ML, Blomqvist C, Ahlgren J, Edlund P, Hansen J, et al. Breast cancer on the Internet: the quality of Swedish breast cancer websites. Breast. 2004;13:376-82.

[20] Quinn EM, Corrigan MA, McHugh SM, Murphy D, O'Mullane J, Hill AD, et al. Breast cancer information on the internet: analysis of accessibility and accuracy. Breast. 2012;21:514-7.

[21] Ream E, Blows E, Scanlon K, Richardson A. An investigation of the quality of breast cancer information provided on the internet by voluntary organisations in Great Britain. Patient Educ Couns. 2009;76:10-5.

[22] Bruce JG, Tucholka JL, Steffens NM, Neuman HB. Quality of online information to support patient decision-making in breast cancer surgery. J Surg Oncol. 2015;112:575-80.

[23] Lynch NP, Lang B, Angelov S, McGarrigle SA, Boyle TJ, Al-Azawi D, et al. Breast reconstruction post mastectomy- Let's Google it. Accessibility, readability and quality of online information. Breast. 2017;32:126-9.

[24] Noar SM, Althouse BM, Ayers JW, Francis DB, Ribisl KM. Cancer information seeking in the digital age: effects of Angelina Jolie's prophylactic mastectomy announcement. Med Decis Making. 2015;35:16-21.

[25]

[26] Eysenbach G, Kohler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. Bmj. 2002;324:573-7.

[27] Chumber S, Huber J, Ghezzi P. A methodology to analyze the quality of health information on the internet: the example of diabetic neuropathy. Diabetes Educ. 2015;41:95-105.

[28] Maki A, Evans R, Ghezzi P. Bad News: Analysis of the Quality of Information on Influenza Prevention Returned by Google in English and Italian. Front Immunol. 2015;6:616.

[29] Yaqub M, Ghezzi P. Adding Dimensions to the Analysis of the Quality of Health Information of Websites Returned by Google: Cluster Analysis Identifies Patterns of Websites According to their Classification and the Type of Intervention Described. Front Public Health. 2015;3:204.

[30] Aslam R, Gibbons D, Ghezzi P. Online information on antioxidants. Information quality indicators, commercial interests and ranking by Google. Frontiers in Public Health. 2017;5:a90.

[31] WebpageFX. Readability Test Tool. 2017.

[32] DuBay WH. The Principles of Readability. Online Submission. 2004.

[33] Reichow B, Halpern JI, Steinhoff TB, Letsinger N, Naples A, Volkmar FR. Characteristics and quality of autism websites. J Autism Dev Disord. 2012;42:1263-74.

[34] Killeen S, Hennessey A, El Hassan Y, Killeen K, Clarke N, Murray K, et al. Gastric cancer-related information on the Internet: incomplete, poorly accessible, and overly commercial. Am J Surg. 2011;201:171-8.

[35] Brosnan S, Barron E, Sahm LJ. Health literacy and the clozapine patient. Perspectives in public health. 2012;132:39-42.

[36] Barnett T, Hoang H, Furlan A. An analysis of the readability characteristics of oral health information literature available to the public in Tasmania, Australia. BMC Oral Health. 2016;16:35.

[37] Wang LW, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: application, results, and recommendations. Res Social Adm Pharm. 2013;9:503-16.

[38] D'Alessandro DM, Kingsley P, Johnson-West J. The readability of pediatric patient education materials on the World Wide Web. Arch Pediatr Adolesc Med. 2001;155:807-12.

[39] Schwitzer G. A guide to reading health care news stories. JAMA Intern Med. 2014;174:1183-6.

[40] Wang MT, Gamble G, Bolland MJ, Grey A. Press releases issued by supplements industry organisations and non-industry organisations in response to publication of clinical research findings: a case-control study. PLoS One. 2014;9:e101533.

[41] Vargas CR, Chuang DJ, Ganor O, Lee BT. Readability of online patient resources for the operative treatment of breast cancer. Surgery. 2014;156:311-8.

[42] Eurostat. Digital economy & society in the EU — a browse through our online world in figures. 2017.

[1] Fox S, Duggan M. Health online 2013. Health. 2013.
[2] Eurostat. Internet access and use statistics - households and individuals. Eurostat Statistics Explained2017.
[3] 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. 1997;277:1244-5.
[4] Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health Websites. Comput Biol Med. 1998;28:603-10.
[5] Dy CJ, Taylor SA, Patel RM, McCarthy MM, Roberts TR, Daluiski A. Does the quality, accuracy, and readability of information about lateral epicondylitis on the internet vary with the search term used? Hand (N Y). 2012;7:420-5.
[6] Fitzsimmons PR, Michael BD, Hulley JL, Scott GO. A readability assessment of online Parkinson's disease information. J R Coll Physicians Edinb. 2010;40:292-6.
[7] Hansberry DR, Agarwal N, Shah R, Schmitt PJ, Baredes S, Setzen M, et al. Analysis of the readability of patient education materials from surgical subspecialties. Laryngoscope. 2014;124:405-12.
[8] Choices N. Breast cancer (female) - Treatment 2016.
[9] Caldon LJ, Walters SJ, Ratcliffe J, Reed MW. What influences clinicians' operative preferences for women with breast cancer? An application of the discrete choice experiment. Eur J Cancer. 2007;43:1662-9.
[10] Caldon LJ, Walters SJ, Reed MW. Changing trends in the decision-making preferences of women with early breast cancer. Br J Surg. 2008;95:312-8.
[11] Stewart MA. Effective physician-patient communication and health outcomes: a review. CMAJ. 1995;152:1423-33.
[12] Fallowfield L, Hall A, Maguire P, Baum M, A'Hern R. A question of choice: results of a prospective 3-year follow-up study of women with breast cancer. The Breast. 1994;3:202-8.
[13] Castleton K, Fong T, Wang-Gillam A, Waqar MA, Jeffe DB, Kehlenbrink L, et al. A survey of Internet utilization among patients with cancer. Support Care Cancer. 2011;19:1183-90.
[14] McLeod J, Yu I, Ingledew PA. Peering into the Deep: Characterizing the Internet Search Patterns of Patients with Gynecologic Cancers. J Cancer Educ. 2017;32:85-90.
[15] Nguyen SK, Ingledew PA. Tangled in the breast cancer web: an evaluation of the usage of web-based information resources by breast cancer patients. J Cancer Educ. 2013;28:662-8.
[16] Taioli E, Joseph GR, Robertson L, Eckstein S, Ragin C. Knowledge and prevention practices before breast cancer diagnosis in a cross-sectional study among survivors: impact on patients' involvement in the decision making process. J Cancer Educ. 2014;29:44-9.
[17] Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. JAMA. 2001;285:2612-21.
[18] Meric F, Bernstam EV, Mirza NQ, Hunt KK, Ames FC, Ross MI, et al. Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites. Bmj. 2002;324:577-81.
[19] Nilsson-Ihrfelt E, Fjällskog ML, Blomqvist C, Ahlgren J, Edlund P, Hansen J, et al. Breast cancer on the Internet: the quality of Swedish breast cancer websites. Breast. 2004;13:376-82.
[20] Quinn EM, Corrigan MA, McHugh SM, Murphy D, O'Mullane J, Hill AD, et al. Breast cancer information on the internet: analysis of accessibility and accuracy. Breast. 2012;21:514-7.
[21] Ream E, Blows E, Scanlon K, Richardson A. An investigation of the quality of breast cancer information provided on the internet by voluntary organisations in Great Britain. Patient Educ Couns. 2009;76:10-5.
[22] Bruce JG, Tucholka JL, Steffens NM, Neuman HB. Quality of online information to support patient decision-making in breast cancer surgery. J Surg Oncol. 2015;112:575-80.
[23] Lynch NP, Lang B, Angelov S, McGarrigle SA, Boyle TJ, Al-Azawi D, et al. Breast reconstruction post mastectomy- Let's Google it. Accessibility, readability and quality of online information. Breast. 2017;32:126-9.
[24] Noar SM, Althouse BM, Ayers JW, Francis DB, Ribisl KM. Cancer information seeking in the digital age: effects of Angelina Jolie's prophylactic mastectomy announcement. Med Decis Making. 2015;35:16-21.