


Quality of Online Resources for Pancreatic Cancer Patients

Lauren De Groot¹ · Ilene Harris² · Glenn Regehr^{3,4} · Ara Tekian² · Paris-Ann Ingledeu^{3,5} 

Published online: 18 October 2017
© American Association for Cancer Education 2017

Abstract The Internet is increasingly a source of information for pancreatic cancer patients. This disease is usually diagnosed at an advanced stage; therefore, timely access to high-quality information is critical. Our purpose is to systematically evaluate the information available to pancreatic cancer patients on the internet. An internet search using the term “pancreatic cancer” was performed, with the meta-search engines “Dogpile”, “Yippy” and “Google”. The top 100 websites returned by the search engines were evaluated using a validated structured rating tool. Inter-rater reliability was evaluated using kappa statistics and results were analyzed using descriptive statistics. Amongst the 100 websites evaluated, etiology/risk factors and symptoms were the most accurately covered (70 and 67% of websites). Prevention, treatment and prognosis were the least accurate sections (55, 55 and 43% of websites). Prevention and prognosis were also the least likely to be covered with 63 and 51 websites covering these, respectively. Only 40% of websites identified an author. Twenty-two percent of websites were at a university reading level. The majority of online information is accurate

but incomplete. Websites may lack information on prognosis. Many websites are outdated and lacked author information, and readability levels are inappropriate. This knowledge can inform the dialogue between healthcare providers and patients.

Keywords Pancreatic cancer · Patient education · Internet · Online resources · Quality

Introduction

The Internet is increasingly a source of health information. Recent estimates suggest that nearly 90% of US adults access the Internet and 72% of Internet users have searched online for health information within the past year [1, 2]. The tendency for patients to search the Internet for specific disease information is increasing, and the majority of online health information seekers start their search with a search engine rather than a site that specializes in health information [3]. Perhaps not surprisingly, cancer is one of the most commonly searched topics and the Internet is one of the most frequently used sources of cancer information [4–6].

Cancer patients access the Internet for a variety of reasons, including to improve their understanding of their diagnosis, to find information about treatments, to learn about living with cancer, and to gain support from others [7]. Gastrointestinal cancer, including pancreatic cancer, is the most commonly searched cancer type on the Internet, comprising 15% of all cancer-related searches [3]. Pancreatic cancer is relatively rare compared with other cancer types, but it has the lowest 5-year survival of all malignancies at 8% [8]. Given the poor prognosis for the disease and the frequency with which these patients use the Internet for health information, access to high-quality information in a timely manner is critical.

✉ Paris-Ann Ingledeu
pingledeu@bccancer.bc.ca

¹ Undergraduate Medical Program, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

² Department of Medical Education, College of Medicine, University of Illinois at Chicago, Chicago, IL, USA

³ Department of Surgery, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

⁴ Center for Health Education Scholarship, University of British Columbia, Vancouver, BC, Canada

⁵ Fraser Valley Cancer Center, British Columbia Cancer Agency, 13750 96th Ave, Surrey, BC V3V 1Z2, Canada

Previous studies have focused on the quality of Internet information for other cancers and gastrointestinal diseases, including colorectal and gastric cancers. Authors of these studies have found variable website quality and accuracy, overrepresentation of commercial websites, and readability scores above the level of the average patient [9–11].

Similar research focusing on pancreatic cancer is limited. Authors of a small study in the UK found that Internet resources for pancreatic and esophageal cancers were of higher quality than for other gastrointestinal cancers, but that overall quality ratings were still low [12]. Storino et al. published the most comprehensive study to date assessing the quality of online pancreatic cancer information [13]. They examined two parameters of patient-oriented online resources for pancreatic cancer: readability and accuracy of treatment specific information. They found that the grade level required for readability of pancreatic cancer treatment information is higher than recommended and that accuracy of information is associated with a more difficult reading level. They also found websites discussing treatment information to be overall accurate, except for those discussing alternative therapies.

In summary, patients with pancreatic cancer are increasingly turning to the Internet to improve their understanding of their disease. While some small studies have examined the quality of currently available websites, there is no research evaluating the quality of online pancreatic cancer information for topic areas other than treatment, such as symptoms, prevention and prognosis. Additionally, there are no previous studies systematically analyzing the completeness, interactivity, accountability and website organization of online pancreatic cancer information. The purpose of this study was to apply a rating tool to systematically evaluate the quality of 100 websites designed to provide pancreatic cancer patient information. The tool was used to evaluate sites with respect to currency, disclosure, attribution, interactivity, completeness and accuracy of content.

Methods

An Internet search using the term “pancreatic cancer” was performed on June 18, 2015, with the meta-search engines “Dogpile”, “Yippy” and the search engine “Google”. The URLs of the first 500 websites for each search engine were recorded. Inclusion and exclusion criteria were applied to the lists to select for websites that were specifically designed for providing patient information and to exclude blogs, primary news or journal articles. Websites had to be available without subscription and sites that provided links to other websites without original content were excluded. Once the lists for each search engine had the inclusion/exclusion criteria applied, the

three lists were combined to provide an average rank order from all three search engines. A list of “the top 100 websites” from all search engines was then compiled.

A structured rating tool, validated and developed by our research group, was applied to evaluate the top 100 websites with respect to attribution/currency, interactivity, readability, content and accuracy [14]. The structured rating tool was developed in 2009 and was adapted based on the Health on the Internet (HON) Foundation code, JAMA, and a detailed review of available resources intended to evaluate the quality of medical information on the Internet [14–16]. The various components of this validated tool have been reported elsewhere. Accountability criteria were derived from the HON code principles [17] and the DISCERN scale [18], an instrument developed at Oxford University to assist people without content expertise to assess the quality of a written health information publications. Interactivity and aesthetic criteria were based on adaptation of the Abbot’s scale [19, 20]. Readability and content quality assessment criteria were established based on the use of several evidence-based resources [20–23]. Readability was assessed using the Flesch-Kincaid (FK) grade level and the SMOG Index. To assess readability, text from the sections on definition, diagnosis and treatments were directly inputted to the readability assessment tool on Read-able.com. Content was evaluated based on the materials deemed by content experts to be informative for a patient’s understanding of pancreatic cancer. To develop a metric to compare for both content and accuracy, the research assistant reviewed pancreatic cancer materials from the National Comprehensive Cancer Network (NCCN) and UpToDate and summarized the information. Two oncologists then reviewed the summary and through iterative discussions developed a consensus document reflecting the essential components required for content and the level of detail required for accuracy levels, ranging from incorrect and mostly correct to completely correct.

The reliability of the website evaluation was determined in two phases. First, the principal investigator and a research assistant used the structured rating tool to independently code a random 20% sample of websites. The sample of websites was determined by assigning a number to each website and using an online random number generator to select the random set of sites. The kappa statistic was used to measure inter-coder reliability [14]. For categories where website rankings showed a kappa value of < 0.70, the two raters met, discussed the rating differences and resolved the discrepancy by consensus. Once consensus was reached, each rater rated a new random sample of 10% of the websites. Following the second phase, kappa values were again defined and any discrepancies discussed. In this study, there were no modifications to the rating tool at this point and the research assistant analyzed the remaining 70% of the websites.

Results

The Internet search yielded 9,810,000 hits on Google and 37,405 hits on Yippy. The search engine Dogpile did not disclose a total number of hits. Over 800 websites were recorded from the three search engines and a list of the top 100 websites covering pancreatic cancer was compiled using pre-determined inclusion and exclusion criteria.

Website Affiliation

Of the top 100 websites, the most common affiliations were non-profit organization (.org) and commercial (.com), which accounted for 43% and 41% of the websites respectively. The remaining website affiliations were academic/university (.edu), other and American government (.gov) with 9, 5 and 2% of websites, respectively.

Accountability

Accountability was examined with respect to the disclosure of authorship, citations, presence of external links and creation modification dates.

Less than half (40) of the 100 websites identified an author, and even fewer identified the author's credentials (34) or author's affiliation (32).

Resources and citations on the sites were evaluated. Reliable sources were defined as journal articles, peer-reviewed sites such as up-to-date, academic or government sites, and textbooks. Fifty-eight of the websites cited at least one reliable source, of which 19 used one reliable source and 39 used two or more reliable sources.

External links were evaluated to ensure current functionality. Fourteen of the 100 websites provided one external link (excluding advertising) and 56 of the websites provided two or more external links. Fifty percent or greater of the links provided were functional in 66 of the websites, whereas the remaining 34 websites either provided no external links (31 sites) or had less than 50% of the links provided that were functional (3 sites).

Eighty-seven of the websites indicated a date of creation. Over half of the websites (61) were last updated over 4 years ago or had no date of last modification identifiable. Seven websites were last updated between 2 and 4 years from the date of analysis, and 32 websites were last updated within 2 years from the date of analysis.

Interactivity

Interactivity was evaluated based on the presence or absence of five features (search engine, audio/visual support, discussion board/forum, queries to webmaster, educational support). The most common interactive tool found was a search engine,

in 93 of the 100 websites analyzed. Twenty-seven websites provided audio/video support and 16 allowed website users to send queries to the webmaster. Thirteen websites contained a discussion board or forum, and eight provided educational support.

Site Organization

The websites were evaluated for five structural tools including headings, subheadings, pictures/diagrams/tables, hyperlinks and absence of advertising. Approximately, a third of websites analyzed used each of five, four and three structural tools in its design (33, 32 and 30, respectively). The remaining five websites used two structural tools.

Readability

Ninety-nine of the 100 top websites were analyzed for readability using the online readability test tool at www.read-able.com. One website was not evaluated as the text could not be directly inputted (i.e. cut and paste) into the readability test tool. Flesch-Kincaid grade level and readability ease and SMOG index were calculated. The majority (73%) of websites were written at a high school reading level (Flesch-Kincaid grades 8.00–12.99). Twenty-two percent of websites were written at a university level (> 13.00), while only 5% were written at an elementary level (< 7.99). The average grade level was 11.7.

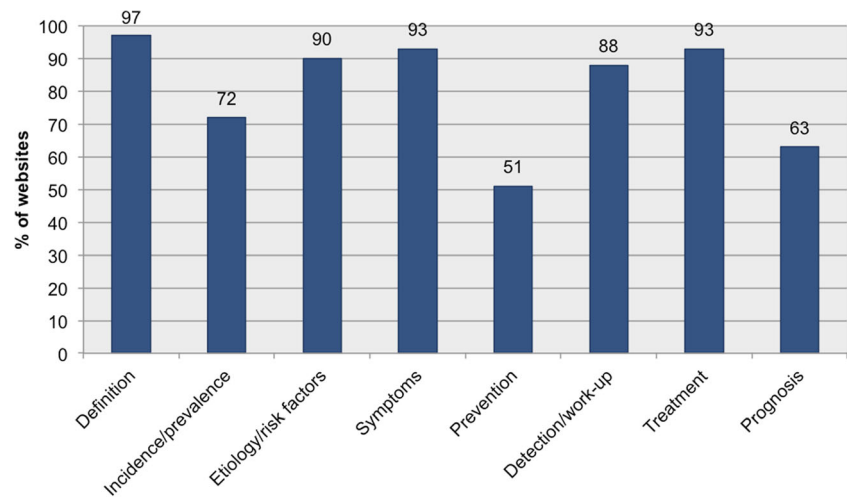
Coverage and Accuracy

Websites were assessed for their coverage of eight pancreatic cancer topics: definition, incidence/prevalence, etiology/risk factors, symptoms, prevention, detection/work-up, treatment and prognosis (Fig. 1). The definition of pancreatic cancer was most consistently covered, followed by treatment (97 and 93 of the 100 websites evaluated). Prognosis and prevention were least likely to be covered, with 63 and 51 websites addressing these two topics, respectively.

The accuracy of each section, if present, was then assessed as “completely accurate and has all required information”, “mostly accurate and/or missing some required information” or “not present or not accurate”. Accuracy and completeness were assessed compared with information from Uptodate.com and the National Cancer Institute.

The information presented on the majority of websites was factually correct. Websites receiving a low score for accuracy/completeness were more often missing important information rather than containing incorrect information (Fig. 2). When the information presented was incorrect, it was often due to out-of-date statistics. Etiology/risk factors and symptoms were most likely to be judged completely accurate and containing all required information (70 and 67% of websites,

Fig. 1 Coverage of pancreatic cancer topics



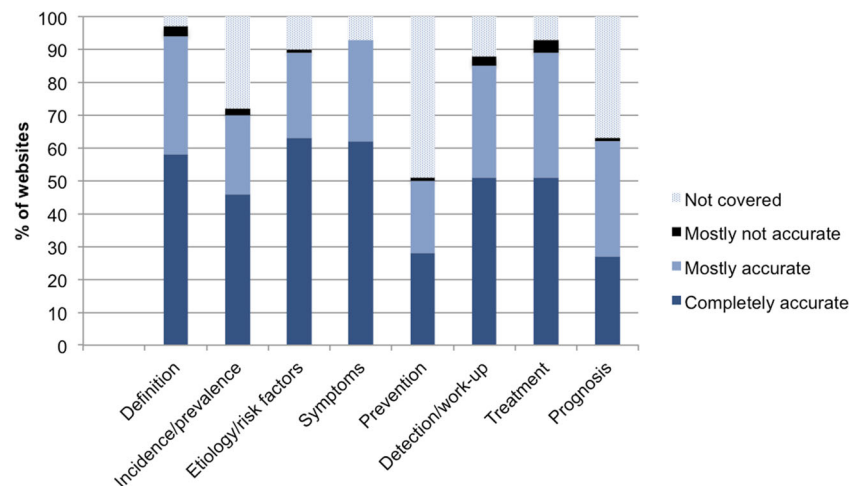
respectively). Prevention, treatment and prognosis were the least accurate sections, judged as completely accurate and containing all required information in 55, 55 and 43% of websites, respectively.

The top 100 websites were evaluated for global accuracy. As indicated above, website content was more often incomplete than factually inaccurate; thus, global accuracy scores were better than the accuracy/completeness scores described above. Eighty-three were judged completely accurate, 17 mostly accurate and 0 mostly not accurate. In terms of objectivity, 87 of the 100 websites expressed no bias.

Overall Quality

Total scores were calculated for each website taking into account authorship, attribution, disclosure, currency, links and their accessibility, interactivity, site organization, readability, coverage, accuracy and objectivity. The maximum possible score was 55. Scores ranged from 51 (www.cancerresearchuk.org) to 12 (www.whipple-procedure.org).

Fig. 2 Accuracy and completeness of pancreatic cancer topics



The average score was 34.2. Table 1 lists the top 10 websites by score.

Discussion

The Internet is an important information source for cancer patients, yet research evaluating the quality of online pancreatic cancer information is limited. This comprehensive study used a validated tool to assess 100 websites for currency, disclosure, attribution, interactivity, content and readability.

Since the majority of patients use a search engine to access health information, rather than direct links from reputable sources, patients require a way to evaluate the quality of the websites retrieved by a search engine [1]. Disclosing authorship and credentials, keeping sites up to date and providing references help patients to make these judgments. These criteria are standard requirements for scientific literature. Less than half of websites evaluated disclosed authorship and even fewer provided the author's credentials or affiliation.

Table 1 Top 10 pancreatic cancer websites based on overall quality score

Website URL	Quality score (maximum 55)
http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/pancreatic-cancer	51
http://www.cancer.org/cancer/pancreaticcancer/	50
http://www.cancer.net/cancer-types/pancreatic-cancer	49
https://www.lustgarten.org/	49
http://www.oncolink.org/types/article.cfm?c=172&id=1739	49
http://www.pancreaticcancer.org.uk/about-us/about-the-charity/about-us-and-our-aims/	48
http://www.cancercouncil.com.au/pancreatic-cancer/what-is-pancreatic-cancer/	48
http://www.cancer.gov/types/pancreatic	47
http://www.cancer.ca/en/cancer-information/cancer-type/pancreatic/pancreatic-cancer	47
http://www.medicinenet.com/pancreatic_cancer/article.htm	47

Only one third of websites had been updated in the past 2 years. Additionally, nearly half of websites cited no reliable sources. These findings that currency, authorship and references are lacking in online pancreatic cancer information parallel studies evaluating websites for gastrointestinal diseases, prostate cancer and colorectal cancer [9, 11, 24]. These shortcomings could make it challenging even for the most knowledgeable patient to get information about their condition online. Educating patients on how to evaluate website quality and credibility, and directing patients to trusted websites, are important solutions.

This study showed that the majority of information provided on the websites with respect to pancreatic cancer was factually correct, although it was common for sites to lack key information. When information was incorrect, it was often because of outdated incidence or prevalence statistics. Studies evaluating the quality of online gastric and prostate cancer information also found websites to be incomplete but generally accurate [10, 24]. As indicated in a study by Black et al., accurate but incomplete information can be misleading for Internet consumers [24]. A significant number of the websites in our study were from US cancer treatment centers targeted towards attracting patients to their facility. These websites lacked links to external information and their treatment sections focused on what was offered at that center. Again, while the information provided was mostly correct, it was incomplete and may prevent patients from becoming fully educated in a non-biased manner. Physicians and cancer treatment centers can provide patients with links to accurate and complete websites to guide their research.

Reading ability is an important component of health literacy. Patient health information should not exceed a 7th grade reading level for consumers to optimally understand the written material [25]. Online information across a range of health topics, including other gastrointestinal cancers, is written well above the optimal grade level for the average health consumer [2, 25]. This study demonstrated that the readability of online

pancreatic cancer information is similarly too high, with only 5% of websites scoring at the appropriate elementary reading level. Using the Flesch-Kincaid grade level readability score, 22% of websites required a university level education to understand. Storino et al. also found inappropriately high readability levels for online pancreatic cancer treatment information, yet found improved accuracy to be associated with a difficult reading level [13]. This finding illustrates that a certain level of technical language may be needed to accurately explain complex diseases. Websites can address this problem by providing glossaries to define terms that may be unfamiliar to patients. Additionally, knowing that the readability levels of most patient websites are too high, providers can ask their patients what kind of information they are finding on the internet, whether it complements or contradicts what they have told them, and help to correct or better explain the information. Storino et al. also recommend the use of visual aids to help the reader understand information when the readability is high [13]. We found some use of such tools to organize website information and improve interactivity, although there remains room for improvement.

Patients who search online for health information may be starting to recognize the shortfalls of the Internet, as described above. Although use of the Internet as a health resource is increasing, the public's trust in online health information has decreased over time [2]. Meanwhile, trust in physicians as a source of health information has increased [2]. These trends may indicate a role for healthcare providers in guiding patients towards high-quality pancreatic cancer resources. Knowing what gaps exist in patient information online, such as prognosis information for pancreatic cancer, can aid in development of better resources and help healthcare providers to address gaps in patient knowledge.

There are several limitations to this study. As indicated by Storino et al., evaluating the completeness of online resources can be misleading as some websites are only designed to cover certain topics [13]. Only English language websites were

included in this study, eliminating resources for patients who seek their information in other languages. For the readability evaluation, representative sections of the websites were evaluated as distinguished from the complete content. It is possible that using the readability tool on all written information on the website may have altered the readability scores; however, it would have been time consuming and unrealistic to have done this for 100 websites. Additionally, all searches were conducted from the same geographic location. It is acknowledged that for some search engines, geographic location can impact the “hits” received. We feel that by utilizing two meta-search engines and one search engine and combining all hits in a systematic manner, we likely minimized the chance that the retrieved websites would be significantly different. Having said that, an area of future interest may be to look at the impact of geographic location on websites retrieved by a search engine for pancreatic and other cancers.

In conclusion, this study systematically applied a validated rating tool to comprehensively evaluate the quality of online pancreatic cancer patient information. This is the largest and most comprehensive review of the literature to date. This study shows that the quality of online pancreatic cancer information is variable. Many websites are outdated and lack author information. While the majority of information presented is factually correct, sites may lack information on prevention and prognosis. The majority of websites are also written at a level too high for the average patient to understand. There is significant room for improvement in the use of structural and organizational tools, including glossaries and visual aids, to offset high readability levels. Healthcare providers can initiate discussions with their patients about the quality of online cancer information and recommend trusted websites to their patients.

Acknowledgements This research was supported by the UBC Summer Student Research Program.

References

- Susannah Fox and Maeve Duggan. 2013. Health online 2013 pew research center. A National survey by the Pew Research Center's Internet & American Life Project <http://www.pewinternet.org/2013/01/15/health-online-2013/> Accessed 6 June 2017
- A Decade of Change in Health Information Technologies and Patient Engagement 2010 https://hints.cancer.gov/docs/Briefs/HINTS_Brief_16.pdf Accessed 6 June 2017
- Bader JL, Theofanos MF (2003) Searching for cancer information on the internet: analyzing natural language search queries. *J Med Internet Res* 5:e31
- Nguyen SKA, Ingledew PA (2013) Tangled in the breast cancer web: an evaluation of the usage of web-based information resources by breast cancer patients. *J Cancer Educ* 28:662–668. <https://doi.org/10.1007/s13187-013-0509-6>
- Hamilton SN, Scali EP, Yu I, Gusnowski E, Ingledew PA (2014) Sifting through it all: characterizing melanoma patients' utilization of the Internet as an information source. *J Cancer Educ* 30:580–584. <https://doi.org/10.1007/s13187-014-0711-1>
- McLeod J, Yu I, Ingledew PA (2017) Peering into the deep: characterizing the Internet search patterns of patients with gynecologic cancers. *J Cancer Educ* 32:85–90. <https://doi.org/10.1007/s13187-016-1002-9>
- Ziebland S (2004) How the internet affects patients' experience of cancer: a qualitative study. *BMJ* 328:564. <https://doi.org/10.1136/bmj.328.7439.564>
- Canadian Cancer Society's Advisory Committee on Cancer Statistics (2015) Canadian cancer statistics 2015. Toronto, ON
- Tangri V, Chande N (2011) Quality of internet-based information on gastrointestinal diseases. *Can J Gastroenterol* 25:93–96
- Killeen S, Hennessey A, El Hassan Y, Killeen K, Clarke N, Murray K, Waldron B (2011) Gastric cancer related information on the Internet: incomplete, poorly accessible, and overly commercial. *Am J Surg* 201. Elsevier Inc.:171–178. <https://doi.org/10.1016/j.amjsurg.2009.12.015>
- Sajid MS, Iftikhar M, Monteiro RS, Miles AFW, Woods WGA, Baig MK (2008) Internet information on colorectal cancer: commercialization and lack of quality control. *Color Dis* 10:352–356. <https://doi.org/10.1111/j.1463-1318.2007.01316.x>
- Chan DSY, Willicombe A, Reid TD, Beaton C, Arnold D, Ward J, Davies JL, Lewis WG (2012) Relative quality of internet-derived gastrointestinal cancer information. *J Cancer Educ* 27:676–679. <https://doi.org/10.1007/s13187-012-0408-2>
- Storino A, Castillo-Angeles M, Watkins AA, Vargas C, Mancias JD, Bullock A, Demirjian A et al (2016) Assessing the accuracy and readability of online health information for patients with pancreatic cancer. *JAMA Surg* 192:841–845. <https://doi.org/10.1001/jamasurg.2016.0730>
- Ingledew, Paris-Ann, Glenn Regehr, Ara Tekian, and Ilene Harris (2015) Navigating the internet in patient physician collaboration (master's thesis). UIC. <http://hdl.handle.net/10027/19489>
- Silberg B (1997) Assessing, controlling, and assuring the of medical information on the internet quality. *JAMA* 277:1244–1245
- Jadad AR, Gagliardi A (1998) Rating health information on the Internet: navigating to knowledge or to Babel? *JAMA* 279:611–614. <https://doi.org/10.1001/jama.279.8.611>
- Health on the Net: HONcode Site Evaluation Form (2017) <https://www.healthonnet.org/HONcode/Conduct.html>. Accessed 6 June 2017
- Chamock D, Shepperd S, Needham G, Gann R (1999) DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 53:105–111
- Abbott VP (2000) Web page quality: can we measure it and what do we find? A report of exploratory findings. *J Public Health Med* 22: 191–197
- Kisely S, Ong G, Takyar A (2003) A survey of the quality of web based information on the treatment of schizophrenia and Attention Deficit Hyperactivity Disorder. *Aust N Z J Psychiatry* 37:85–91
- Griffiths KM, Christensen H (2002) The quality and accessibility of Australian depression sites on the World Wide Web. *Med J Aust* 176S:S97–104
- Griffiths KM, Christensen H (2000) Information in practice depression: cross sectional survey. *BMJ* 321:1511–1515
- Williamson JML, Hocken DB (2011) Assessing the internet prevalence of cancer. *Int J Clin Pract* 65:178–181
- Black PC, Penson DF (2006) Prostate cancer on the internet—information or misinformation? *J Urol* 175:1836–1842. [https://doi.org/10.1016/S0022-5347\(05\)00996-1](https://doi.org/10.1016/S0022-5347(05)00996-1)
- Walsh TM, Volsko TA (2008) Readability assessment of internet-based consumer health information. *Respir Care* 53:1310–1315