Dr Google - Assessing the reliability and readability of information on General Surgical Procedures found via search engines

Raziqah Ramli[1,2] MD, MSc,

Maxwell Andrew Jambor[1] MD,

Chia Yew Kong [2, 3] MB ChB, MSc, MRCS

1- School of Medicine, University of New South Wales, Australia

2 - School of Clinical Sciences, University of Edinburgh, Scotland

3 - School of Medicine, University of Glasgow, Scotland

Correspondence:

Raziqah Ramli

Raz.ramli9@gmail.com

Department of Surgery, The Canberra Hospital, Yamba Drive, Garran ACT 2605, Australia

+61450129993

The corresponding author is not a recipient of a research scholarship. This paper has not been submitted elsewhere for publication.

Email: raz.ramli9@gmail.com

Keywords: patient information, surgical procedures, readability, reliability, search engine, appendicectomy, cholecystectomy, hernia repair, colectomy, haemorrhoidectomy

Funding/Conflicts of interest: There was no specific funding of this particular study. All authors declare that they do not have any conflicts of interest to disclose.

Number of figures and tables included: 1 Figure, 7 Tables

Abstract

Background

The most common general surgical emergency operations are laparoscopic appendicectomy, laparoscopic cholecystectomy, hernia repair, hemorrhoidectomy and colectomy. Patients commonly perform an internet search for more information prior to undergoing surgery, which can lead to an inappropriate understanding of their procedure. The aim is to assess the quality of information available on three of the most used search engines.

Methods

A search was conducted on Google.com, Bing.com and Yahoo.com using the terms related to laparoscopic appendicectomy, laparoscopic cholecystectomy, hemorrhoidectomy, hernia repair and colectomy. First 20 results from each search engine were collected for evaluation. Results were excluded if they were sponsored, duplicates, academic publications, advertisements, forums, audiovisual tools, social media, or any non-English information. Included results were assessed for reliability using DISCERN and JAMA benchmark score. Readability was assessed using Flesch Reading Ease (FRE) Score and Simple Measure of Gobbledygook (SMOG).

Results

197 websites were analysed, 44.7% were published by institutions, 34.5% by health websites and 20.8% by independent surgeons. Mean DISCERN scores for Institutions was 54.6±11.3, independent surgeons 45.9±11.4 and health websites 58.7±10.3. Mean JAMA score for Institutions was 1.0±1.0, independent surgeons 0.1±0.4 and health websites 1.7±1.1. FRE scores for institutions was 51.6±10.3, independent surgeons 40.9 ± 10.2, and health websites

45.7±12.3. SMOG scores were 9.8±1.5 for institutions, 11.4±1.6 for independent surgeons and 10.6±1.7 for health websites.

Conclusion

Health information on common general surgical procedures found on search engines are generally fair to good quality but still above the suggested reading level of the population. Information on surgical procedures should be written at recommended reading level of 13-14 years old.

Keywords

patient information, surgical procedures, readability, reliability, search engine, appendicectomy, cholecystectomy, hernia repair, colectomy, haemorrhoidectomy

Introduction

Based on the database of Australian Institute of Health and Welfare, general surgical procedures make up approximately one-fifth of hospital admissions for the last five to six years [1]. The most common surgical procedures reported for emergency admissions in the year 2019-2020 include laparoscopic appendicectomy, laparoscopic cholecystectomy, hernia repair, hemorrhoidectomy and colectomy [1]. Patients commonly perform an internet search to seek more information on their procedure before undergoing it[2].

However, these internet searches are often associated with inappropriate or incorrect patient understanding of their procedure [4]. Despite spending time to explain information to patients regarding a procedure, only 40% of the information given was recalled correctly during consultation [5]. It has been reported that physicians spend more time with patients debunking misinformation found online, thus decreasing their efficiency during consultation.[6]

Despite evidence showing that the use of internet has improved patient participation in decision making, patients can be led into perceiving incorrect information available on the internet as valid and accurate. This is a potential danger of using search engines to yield health information as the quality of information on the internet is frequently unregulated [7]. Information may also not be presented in a way that is easily readable or understood by the general public [7], and can result in confusion, being overwhelmed and develop inappropriate expectations [6, 8]. This is important as higher levels of patient satisfaction are associated with pre-operative patient understanding and knowledge about their procedure and condition,

along with improved compliance, health outcomes and reduced healthcare costs in the long term [4, 9].

There are multiple studies investigating the readability and accuracy of health information and websites found on search engines, regarding individual procedures and specific diseases. However, there are no current studies available that has investigated and analysed the quality of health information on the common general surgical procedures that are commonly performed into a single study for comparison.

Therefore, the aim of the study is to analyse the reliability and readability of health information found on popular search engines about the top five most common general surgical procedures.

Methods

A search was conducted by two independent reviewers in the same day on Google.com, Bing.com and Yahoo.com, on 18th of December 2021 in Australia, using the following terms "appendicectomy" or "appendix surgery, "cholecystectomy" or "gallbladder surgery", "hemorrhoidectomy" or "haemorrhoid surgery", "hernia surgery" or "hernia repair", "colectomy" or "colon surgery". The most preferred search engine of choice in Australia is Google, and together with Yahoo and Bing, make up for 98% of market share at the time of search[3]. The search was performed under "incognito mode" or "private browsing" with a deleted search history on a public computer to reduce the chances of biased and tailored results based on previous search terms. VPN was not used during the search.

The first 20 results for each term from the various search engines were included for evaluation, as 95% of patients looked at 15 websites or less [10] and another study found that patients seldom looked past the first page of results [11].

Results were excluded if they included irrelevant or inappropriate content, commercial only websites, links to scientific articles of abstracts, duplicate websites, forums, social media content, videos, online medical dictionaries, websites with broken links or any non-English information[12].

Included websites were categorised into "Health websites", "Institutions", "Independent Surgeons". Articles categorized into Health websites were published by a non-official source, such as for-profit or non-profit companies producing health content not affiliated to

an academic institution, hospitals, or government bodies. Articles included in the "Institution" category had information published by an academic institution, university, hospital, government body. Information published by an independent medical practitioner, or a group of medical practitioners not related to a hospital or university or government body, was classified as "Independent Surgeons".

Two independent reviewers from a medical background analysed the reliability and readability of the articles with the following rating tools.

Rating tools

DISCERN

DISCERN is a standardized tool developed by the Division of Public Health and Primary Care at Oxford University, to assess the content quality of consumer health information [2, 13]. It has 15 questions divided in three sections to evaluate reliability, quality of content and an overall impression, with scores given to the respective sections. It can be used to assess the quality of information without the need for specialist knowledge by looking into whether the sources of evidence were clearly stated, if information is biased and fails to mention a range of options for treatment [13]. The higher the score, the better the quality of the information. The scores can be interpreted as follows: 63 to 80 points = excellent; 51 to 62 = good; 39 to 50 = fair; 27 to 38 = poor; 15 to 26 = very poor[13].

Journal of the American Medical Association (JAMA)

The Journal of the American Medical Association (JAMA) benchmark score consists of 4 components: 1 point for disclosure of authorship, 1 point for attributions of sources, 1 point for disclosure of conflict of interest and 1 point for currency of information[14]. A maximum score of 4 can be achieved. It is known to correlate with higher levels of accuracy and a relatively easy tool to use to assess reliability [14].

Flesch-Kincaid Reading Ease (FRE)

Flesch-Kincaid Reading Ease (FRE) score is used in most readability studies [15] and is calculated using a formula to calculate readability based on the average sentence length and the average number of syllables per word[16]. It has a high retest and inter-rater reliability [17].

The score calculated will range from 0 to 100; the higher the score, the easier the information is to read (Table 1). Low scores indicate the text being more difficult to read.

Simple Measure of Gobbledygook (SMOG)

The Simple Measure of Gobbledygook (SMOG) is a formula to estimate the number of years of education (based on American schooling grade system) an individual needs to understand the article (Table 2)[18]. It is calculated based on a formula that is derived from the square root of the total number of syllables in 30 selected sentences, and adding 3 to the approximate square root [18]. It has been proven to be more valid amongst other readability formulas [18].

Data analysis

Statistical analysis was performed using a statistical software (SPSS). A p value of less than 0.05 is deemed significant. Normality tests were performed to assess the distribution of the data.

Results

600 results were yielded from the first 20 results of each search term from all three search engines. After removing duplicates and screening results using the exclusion and inclusion criteria, 197 results were analysed using the scoring tools by two independent raters (Figure 1). Out of 197 results, 44.7% were published by institutions, 34.5% by health websites and 20.8% by surgeons (Table 3). The majority of the information was published by sources from USA (Table 4). Normality of data was assessed and seen to be parametric.

For reliability, the overall mean DISCERN scores for Institutions was 54.6 ± 11.3, for independent surgeons was 45.9 ± 11.4 and for health websites 58.7 ± 10.3 (Table 4). The mean JAMA score for Institutions was 1.0 ± 1.0, for independent surgeons was 0.1 ± 0.4 and for health website was 1.7 ± 1.1.

For readability, the overall FRE scores from institutions was 51.6 ± 10.3, for independent surgeons was 40.9 ± 10.2 and for health websites 45.7± 12.3 (Table 4). The mean grade level calculated by SMOG grade level scores were 9.8 ± 1.5 for institutions, 11.4 ± 1.6 for independent surgeons and 10.6 ± 1.7 for health websites.

There is a statistically significant difference in all 4 scoring systems between the three website sources for both Reliability and Readability as determined by one-way ANOVA as seen in Table 5.

Reliability and Readability based on procedures

Laparoscopic Appendicectomy

For laparoscopic appendicectomy, health websites have the highest mean DISCERN rating with a score of 60.2 ± 8.3, indicating 'good' quality of information and the highest JAMA score of 1.9 ± 0.7. Independent Surgeons have the lowest mean DISCERN score of 43.2 ± 13.1 and JAMA score of 0.4 ± 0.7. FRE score was the lowest at 36.3 ± 13.0 and SMOG grade level of 11.6 ± 1.6.

Laparoscopic cholecystectomy

For laparoscopic cholecystectomy, DISCERN scores were the highest for health website scoring 60.2 ± 8.3 indicating 'good' quality and independent surgeons having the lowest score of 43.2 ± 13.1, indicating 'fair quality'. JAMA scores were again highest for health websites (1.9 ± 0.7) and lowest for independent surgeons (0.4 ± 0.7). FRE scores were highest at 53.8 ± 10.3 for institutions, and lowest at 36.3 ± 13.0, translating to 'difficult' for independent surgeons. SMOG grade level scores were lowest for institutions 9.3 ± 1.4 and highest for independent surgeons 11.6 ± 1.6.

Colectomy

DISCERN scores were highest with a score 61.4 ± 9.0 for health websites and again lowest for independent surgeons at 37.9 ± 8.1. JAMA scores were highest (2.0 ± 0.9) for health websites, and repeatedly the lowest for Independent Surgeons (0.0 ± 0.0). FRE scores were highest for institutions (53.0 ± 10.0) and lowest for health websites (47.2 ± 15.6). SMOG grade level scores were lowest in the institution group (9.4 ± 1.5) and highest for independent surgeons (10.6 ± 2.2).

Hernia Repair

Health websites had the highest DISCERN and JAMA scores of 58.7 ± 10.3 (‘good’ quality) and 1.7 ± 1.3 respectively, in contrast to independent surgeons with the lowest DISCERN and JAMA score 48.4 ± 8.6 (‘fair’ quality) and 0.0 ± 0.0 respectively. Institutions had the most readable information on hernia repair with the highest FRE scores at 47.4 ± 9.2 (‘difficult’ level) and the lowest SMOG grade level of 10.6 ± 1.8. Independent surgeons had the lowest readability with FRE score of 39.6 ± 9.5 (‘difficult’) and SMOG grade levels of 11.9 ± 1.6.

Hemorrhoidectomy

For hemorrhoidectomy, DISCERN and JAMA scores were highest for health websites (59.0 ± 11.8, 1.7 ± 1.3 respectively), with independent surgeons having the lowest DISCERN and JAMA scores (48.4 ± 8.6, 0.0 ± 0.0 respectively),. FRE scores were highest at 47.4 ± 9.2 for institutions and lowest at 39.6 ± 9.5 for independent surgeons. SMOG grade levels were lowest for health websites at 10.6 ± 1.8 and highest for independent surgeons at 11.9 ± 1.6.

One-way ANOVA between procedures

A one-way ANOVA was also performed to look at the differences in scores between procedures, which showed that only JAMA, FRE and SMOG scores were significantly different between procedures, f=4.69 p=0.00, f=3.53 p=0.01, f=7.01 p=0.00, respectively (Table 6).

Correlations between reliability and readability

Pearson correlation showed coefficient correlation of 0.50, p=0.00 between DISCERN and JAMA scores, and -0.91, p= 0.00 between FRE and SMOG and correlation of -0.12, p=0.02 between JAMA and SMOG scores.

Inter-rater correlation

The inter-rater correlation coefficient was 0.599 (p=0.00), which indicates good inter-rater correlation for scoring the articles.

Discussion

Limited health literacy has been associated with worse outcomes and higher mortality rates [19]. Health literacy is closely tied with literacy levels, as being able to read and comprehend health information influences an individual to make healthcare decisions that can maintain and improve quality of life [20]. In Australia, a reading level of year 8 or equivalent to 13-14 years old is required for comprehension of information across the population[21]. In the US, the current recommendation made by the American Medical Association states that health information is to be written at or below sixth grade reading level or equivalent to 11-12 years old [19].

The readability of health information across all sources however is still “fairly difficult” based on FRE scores and requires a minimum school level of Year 10-12, equivalent to 15-18 years old to be understood. This is above the suggested level for the population to appropriately comprehend and concurs with previous studies [20, 21]. Despite many studies

having proven and reiterating this, many sources still publish information that is too difficult to read for the population, thus making groups with low literacy continue to struggle in comprehending the information currently available [22].

Based on the overall results, the reliability quality of health information ranges from fair to good quality based on DISCERN scores. However, average JAMA scores across all procedures and all website sources were two or less, which means a large proportion of websites struggle to meet even half of the JAMA criteria. This is similar to a previous study by Alshaikh et al. (2021), demonstrating the average websites only met one JAMA criteria [8].

Information provided by health websites powered by profit or non-profit companies have consistently achieved highest reliability scores according to DISCERN criteria for all procedures, amongst the other sources. They provided a wide range of information from information on the condition, what would happen if no treatment was done, benefits and risks of the procedure as well as alternative options available. The author, accreditations, and attributes, with the date of publication to assess for currency were also clearly listed, demonstrated by achieving the highest score for JAMA criteria across all surgical procedures. In addition, a qualitative observation that was made in the study was that information published by health websites was spaced out and there was ample use of dot points and small paragraphs, making it easier and less overwhelming to read. Information published by profit or non-profit companies often employ editorial media teams to produce content that aligns with the DISCERN principles and JAMA criteria [23], for their information to be credible and create trust amongst internet users.

Information published by institutions were some the easiest to read amongst other sources for most of the procedures, with highest FRE scores and the lowest SMOG scores. On the contrary, they have lower DISCERN and JAMA scores when compared to health websites despite having a similar range of information to them that is easier to read. It is common to have a bias that institutions, especially those with popular academic reputations to assume that the information published would be more reliable [9], but this study has proven that they are inferior to health websites. An ANOVA analysis reveals that there was a significant difference in reliability scores between the three sources. However, they still scored higher in reliability than independent surgeons.

Health information published by surgeons has the lowest overall DISCERN, JAMA and FRE scores, and the highest SMOG scores. The average JAMA scores for surgeons were less than one, indicating that they did not fulfil the JAMA criteria for reliability. Reliability is rated as the lowest amongst the other sources of information and the hardest to read, despite having specialist training, accreditations, and qualifications. It was also found that most of the information published only described the procedures and post-operative care, with the occasional explanation of complications in minimal detail. There was no date of publication to know if the information was current. Information was presented in bulky paragraphs, making it more difficult to follow through. It has been known that patients have discounted high-quality information due to poor website design [7].

This is particularly concerning as medical practitioners may not be aware in following DISCERN principles or JAMA criteria when publishing information and mainly rely on their authority and status to warrant the reliability of the content [23]. They may also foster beliefs

that adding unnecessary information and content may confuse the patients instead, however not realizing that patients are now more willing to be well informed about their condition and procedure [23]. Having information that has low reliability and readability may lead to patients distrust with medical practitioners and misinterpreting information may lead to inappropriate healthcare decisions [20]. Given the incidence of rising misinformation on health conditions, it is important for medical practitioners to publish health information that is reliable and readable for the mass population.

All sources, especially independent surgeons need to review the information available on their website and make improvements to produce patient-oriented material using the DISCERN and JAMA criteria to increase reliability, as none of sources have scored 'Excellent' quality. It is acknowledged that institutions and companies do not use a consistent criteria and may base their development on a specific target audience [9], but a committee could be formed to include both medical practitioners and lay people to produce health information of a higher quality[2]. All sources should ensure that the information published should be revised to a readability of level of at least 13-14 years old as previously recommended [21].

There were a few limitations to this study. Data collection could have included patients and lay people rating the website to reduce any medical biases and gauge their perspective from a patients' point of view. The assessment should not only be limited to evaluate written information but also assess the visual aspects of the website such as layout and the use of diagrams and videos [20], using validated tools available as observations have been made during data collection that some websites were visually appealing, which was not reflected in the readability scores [24].

A comparison and correlation of the available results with a “gold standard” source of information would be ideal. However, there is currently no “gold standard” website or article available and it is difficult to determine how a website that could achieve a high DISCERN, JAMA, FRE and low SMOG grade level scores. Further analysis that could be performed in future studies is the assessment of reliability and readability of websites based on the order of search engine results to assess the information that internet users most frequency access.

Conclusion

Health information on the five most common general surgical procedures found on search engines are generally fair to good quality in terms of reliability but overall, still above the suggested reading level of the population. Surgeons need to improve on the quality of information published. Information on surgical procedures should be written at recommended reading level of 13-14 years old.

Acknowledgements

This research was undertaken as part fulfilment of the MSc in Surgical Sciences at the University of Edinburgh.

Disclosure statement

There was no specific funding of this study. All authors declare that they do not have any conflicts of interest to disclose.

References

1. Admitted patient care 2019–20 - Australian hospital statistics. 2019.
2. Musbahi A, Ali N, Brown L, Brown S, Viswanath YKS, Etherson K, et al. A Systematic Review of Online Patient Resources to Support Shared Decision Making for Laparoscopic Cholecystectomy. World journal of surgery. 2021;45(9):2719-33.
3. Benefits from more competition in internet search [press release]. Australia: Australian Competition & Consumer Comission, 28 October 2021 2021.
4. Danquah G, Mittal V, Solh M, Kolachalam RB. Effect of Internet use on patient's surgical outcomes. Int Surg. 2007;92(6):339-43.
5. Bravo BN, Postigo JML, Segura LR, Selva JPS, Trives JJR, Córcoles MJA, et al. Effect of the evaluation of recall on the rate of information recalled by patients in Primary Care. Patient Education and Counseling. 2010;81(2):272-4.
6. Hungerford D. Internet Access Produces Misinformed Patients: Managing the Confusion. Orthopedics. 2009;32(9).
7. Dawson JQ, Davies JM, Ingledew P-A. Quality of Online Information Regarding Cervical Cancer. Cureus. 2020;12(8):e9511.
8. Alshaikh L, Shimozono Y, Dankert JF, Ubillus H, Kennedy JG. Evaluation of the Quality and Readability of Online Sources on the Diagnosis and Management of Osteochondral Lesions of the Ankle. Cartilage. 2021:19476035211021910.
9. Daraz L, Morrow AS, Ponce OJ, Beuschel B, Farah MH, Katabi A, et al. Can Patients Trust Online Health Information? A Meta-narrative Systematic Review Addressing the Quality of Health Information on the Internet. Journal of General Internal Medicine. 2019;34(9):1884-91.
10. Chang K, Berthelet E, Grubbs E, Hamilton S, Karvat A, Tran E, et al. Websites, Websites Everywhere: How Thyroid Cancer Patients Use the Internet. Journal of cancer education : the official journal of the American Association for Cancer Education. 2020;35(6):1177-83.
11. Tamhankar AP, Mazari FA, Everitt NJ, Ravi K. Use of the internet by patients undergoing elective hernia repair or cholecystectomy. Ann R Coll Surg Engl. 2009;91(6):460-3.
12. Lorenzo-Pouso AI, Pérez-Sayáns M, Kujan O, Castelo-Baz P, Chamorro-Petronacci C, García-García A, et al. Patient-centered web-based information on oral lichen planus: Quality and readability. Med Oral Patol Oral Cir Bucal. 2019;24(4):e461-e7.
13. DISCERN. The DISCERN Instrument. 1997.
14. Chumber S, Huber J, Ghezzi P. A Methodology to Analyze the Quality of Health Information on the Internet: The Example of Diabetic Neuropathy. The Diabetes Educator. 2014;41(1):95-105.
15. Phan A, Jubril A, Menga E, Mesfin A. Readability of the Most Commonly Accessed Online Patient Education Materials Pertaining to Surgical Treatments of the Spine. World Neurosurg. 2021;152:e583-e8.
16. Huynh MNQ, Hicks KE, Malic C. Assessment of the Readability, Adequacy, and Suitability of Online Patient Education Resources for Benign Vascular Tumours Using the DISCERN Instrument. Plastic surgery (Oakville, Ont). 2019;27(4):325-33.

17. Thomas G, Hartley RD, Kincaid JP. Test-Retest and Inter-Analyst Reliability of the Automated Readability Index, Flesch Reading Ease Score, and the Fog Count. Journal of Reading Behavior. 1975;7(2):149-54.
18. Mc Laughlin GH. SMOG Grading-a New Readability Formula. Journal of Reading. 1969;12(8):639-46.
19. Miles RC, Baird GL, Choi P, Falomo E, Dibble EH, Garg M. Readability of Online Patient Educational Materials Related to Breast Lesions Requiring Surgery. Radiology. 2019;291(1):112-8.
20. Cheng C, Dunn M. Health literacy and the Internet: a study on the readability of Australian online health information. Australian and New Zealand Journal of Public Health. 2015;39(4):309-14.
21. Ferguson C, Merga M, Winn S. Communications in the time of a pandemic: the readability of documents for public consumption. Australian and New Zealand Journal of Public Health. 2021;45(2):116-21.
22. Storino A, Castillo-Angeles M, Watkins AA, Vargas C, Mancias JD, Bullock A, et al. Assessing the Accuracy and Readability of Online Health Information for Patients With Pancreatic Cancer. JAMA surgery. 2016;151(9):831-7.
23. Ng MW, Smith R, Wickramesinghe N, Smart PJ, Lawrentschuk N. Health on the Net: Do Website Searches Return Reliable Health Information on Hemorrhoids and Their Treatment? Int Surg. 2017;102(5-6):216-21.
24. McInnes N, Haglund BJ. Readability of online health information: implications for health literacy. Inform Health Soc Care. 2011;36(4):173-89.

Figure legends

Figure 1 – Flowchart showing the process of inclusion of websites for analysis