Quality and Readability of Google Search Information on HoLEP for Benign Prostate Hyperplasia

Yam Ting Ho 1,*, Jeremy Saad 1,2, Femi E. Ayeni 2, Sachinka Ranasinghe 1, Mohan Arianayagam 1, Bertram Canagasingham 1, Ahmed Goolam 1, Nicola Jeffery 1, Mohamed Khadra 1,2, Raymond Ko 1,2, Nicholas Mehan 1, Celi Varol 1, Jonathan Kam 1 and Isaac A. Thangasamy 1,2

1 Nepean Urology Research Group, Nepean Hospital, Kingswood, NSW 2747, Australia
2 Faculty of Medicine, University of Sydney, Sydney, NSW 2050, Australia; femi.ayeni@sydney.edu.au
* Correspondence: jeremy.yt.ho@gmail.com

Abstract: Objective: To assess the quality and readability of online information on holmium laser enucleation of the prostate in managing benign prostate hyperplasia using the most-used search engine worldwide, Google. Methods: Google search terms “Holmium laser surgery” and “enlarged prostate” were used to generate 150 search results. Two independent authors (i) excluded any paywall, scientific literature, or advertisement and (ii) conducted an independent assessment on information quality, which was based on DISCERN, QUEST, and JAMA criteria, and readability, which was based on the FKG, GFI, SMOG, and FRE scores on qualified webpages. A third author was involved if there were any discrepancies between the assessments. Results: 107 qualified webpages were included in the data analysis. The median DISCERN score was 42 out of 80 (IQR 35–49). The median JAMA score was 0 out of 4 (IQR 0–1). The median QUEST score was 9 out of 28 (IQR 9–12). Using the non-parametric ANOVA and post hoc Games–Howell test, significant differences were identified between rankings of webpages. Sponsorship had no influence on the quality of webpages. The overall readability level required a minimum reading level of grade 11. Linear regression analysis showed that a higher ranked webpage is a positive predictor for all three quality assessment tools. Conclusions: The overall quality of online information on HoLEP is poor. We identify that the top-ranked google searches have a higher DISCERN score and are a positive predictor for DISCERN/QUEST/JAMA. Quality online information can benefit patients but should be used in conjunction with professional medical consultation.

Keywords: urological surgical procedures; prostatectomy; transurethral resection of prostate; prostatic hyperplasia; consumer health information; health literacy; health education

1. Introduction

Lower urinary tract symptoms (LUTS) are common in men over 50 years of age. They pose a significant burden of disease and impact quality of life (QoL) [1,2]. In males, LUTS are commonly caused by benign prostatic obstruction (BPO). BPO is a result of the unregulated proliferation of prostatic tissue—this histological diagnosis is known as benign prostate hyperplasia (BPH). The continuous proliferation leads to benign prostate enlargement (BPE), which ultimately can cause obstruction at the bladder outlet [2].

Transurethral resection of the prostate (TURP) has been the gold standard for surgical management of BPH. However, there has been a steady decline in this procedure globally due to the introduction of laser prostatectomy procedures for the resection of larger prostates [3,4] as they incur less complications compared to TURP. It is observed that, since 2009, Australia has been trending towards minimally invasive options when treating BPH [5]. For instance, Holmium laser enucleation of the prostate (HoLEP), an example of a laser prostatectomy procedure, accounted for 5.9% of all BPH procedures in Australia in
and accounted for roughly 9.2% of all BPH procedures in 2022, according to the recent medicare statistics report (see Table S1) [7].

Current EUA and AUA guidelines consider HoLEP as a size-independent treatment option for BPH, and it can be offered as an alternative to TURP for patients with moderate to severe LUTS. Often, HoLEP is used for larger prostates (>80 mL) [2], and it is comparable to TURP in terms of mid- to long-term efficacy (especially for small- to mid-sized prostates). It also shares similar peri-operative outcomes and complication rates at 2nd and 3rd year follow ups [2,8,9]. Furthermore, HoLEP is associated with reduced bleeding risk, shorter hospital stays, reduced reoperation rate, and can be offered to patients who remain on anti-platelet medications [5,6,9].

Online health information has become a vital part in patients' decision-making in health management. Evidence suggests that there is an increasing number of patients using online health information as a tool for further education [10]. It is found that patients accessing online health information are generally from a higher socioeconomic background, while the amount of usage or access to online information is inversely related to their age [10]. While there are several studies assessing the quality of online health information related to various urological topics, such as andrology, urological cancers, and robotic surgery [11–13], none of them focus on HoLEP specifically. Hence, the aim of this study is to assess the quality and readability of online health information related to HoLEP on men with BPH. It is noted that the study from Chen et al. in 2014 was able to utilize the health on the net (HON) tool to evaluate the online health information on various BPH procedures [14]. Unfortunately, the HON tool service has been discontinued; hence, it has not been selected as a quality assessment tool in this study.

2. Methods

2.1. Webpage Selection

Google searches were conducted for the term “Holmium laser surgery” and “enlarged prostate” on the 8 July 2023. Google Chrome Version 114.0.5735.198 was used, and all previous cookie history was deleted to prevent results from being affected by tracking cookies. The first 150 webpage results were reviewed based on inclusion/exclusion criteria. The inclusion criteria are (i) English webpages only and (ii) webpages including information on HoLEP. The exclusion criteria are (i) advertisements, (ii) webpages that were behind a paywall, and (iii) webpages that were presented as scientific literature, such as journal articles or books. Each webpage was scored by two independent reviewers using quality and readability measures. If there was a discrepancy in the score, a third reviewer decided on the final score.

2.2. Quality Measures

To assess each webpage, three recognized quality assessment tools were used: (i) DISCERN, (ii) QUEST, and (iii) the Journal of the American Medical association (JAMA) criteria. DISCERN is a validated questionnaire to assess consumer health information without specialist knowledge. It focuses on areas such as reliability, currency, balance, source, and the quality of information. There are 16 questions and each are assessed against a 5-point Likert-type scale with a score range from 16 to 80. Questions 1–8 are under subsection 1 addressing reliability; questions 9–15 are under subsection 2 addressing details on treatment choices; and question 16 is under subsection 3 measuring the overall quality of the text [15]. QUEST focuses on areas such as authorship, attribution, conflicts of interests, currency, complementarity, and tone. It includes six questions that are weighted and assessed against a Likert-type scale with a score range from 0 to 28 [16]. JAMA criteria is used to assess whether health information is both authentic and reliable. It focuses on authorship, attribution, disclosure, and currency with a score range from 0 to 4 [17].
2.3. Readability Measures

To assess the readability of each webpage, four validated measures were used: (i) the Flesch–Kincaid Grade Level (FKG); (ii) Gunning–Fog Index (GFI); (iii) Simple Measure of Gobbledygook (SMOG); and (iv) Flesch Reading ease score (FRE). Readable, an online readability score calculator, was used to calculate the four readability scores for each webpage. This was achieved by copying the body text from each webpage into the score text function of Readable.

FKG helps to identify the USA school grade level of education needed to understand a text. The higher the score which ranges from 0 to 18, the higher level of education that is needed to understand the text. FRE is similar to FKG. While they are both derived from a similar formula, FRE has different weighting factors and the score ranges from 0 to 100. GFI closely resembles FKG; however, it focuses on complex words. The higher the score which ranges from 0 to 20, the higher level of education that is needed to understand the text. The SMOG index is a tool used to measure the level of comprehension and the number of years of education an average person needs to understand a text [18].

In summary, these scores are equivalent to the USA school grade level of comprehension. For the general public, a text should aim for a score equivalent to the USA school grade level 8, e.g., FKG 8, FRE 70–80, GFI 8, and a SMOG below 10 [18,19].

2.4. Statistical Analysis

JASP 0.17.3 (intel) statistical software was used. Descriptive data were reported as the median and interquartile range. Kendall tau correlation was used to assess the correlation between quality and readability measures. A non-parametric Kruskal Wallis test was used to compare the effects of rank and sponsorship on quality assessment scores. Any significant difference identified was then assessed with a post hoc Games–Howell test. A linear regression model, using rank and sponsorship to predict quality assessment scores, was conducted. Confounds or covariates were excluded in a step-by-step approach to identify significant regression coefficients.

3. Results

Among the 150 reviewed webpages, 107 qualified webpages were included as part of the data analysis and the remaining 43 webpages were excluded. The most common targeted audience were patients (98%) and the remaining were health professionals. The most common sponsorship of these qualified webpages are from urologists' private practice website—with a total of 40 (37%) webpages. Other sources included hospital, 28 (26%); health organization, 15 (14%); university, 10 (9%); non-profit organization, 9 (8%); media/news, 3 (2.8%); and others, 2 (1.8%).

For the purpose of rank allocation, the first 50 webpages shown were considered as top ranked, while the subsequent 50 webpages and the last 50 webpages were considered as middle and bottom ranked, respectively. Among the 107 qualified webpages, 46 were top ranked, 31 were middle ranked, and 30 were bottom ranked.

3.1. Descriptive Statistics

DISCERN scores: the median subsection 1 score was 22 (IQR 19–25); the median subsection 2 score was 18 (IQR 14–21); and the median subsection 3 score was 2 (IQR 1–2). Top ranked webpages had higher median DISCERN scores compared to bottom-ranked webpages (Table 1). Readability: GFI median score was 12.8 (IQR 11.8–14); FKG median score was 10.7 (IQR 9.6–11.7); SMOG median score was 12.8 (IQR 12.1–13.7); and FRE median score was 44.9 (IQR 40.1–51.2).
Table 1. The median total scores (IQR/interquartile range) of all quality assessment tools, along with the median scores (IQR) based on webpages’ rank allocation.

<table>
<thead>
<tr>
<th></th>
<th>Total Score</th>
<th>Top Ranked</th>
<th>Middle Ranked</th>
<th>Bottom Ranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCERN</td>
<td>42 (35–49)</td>
<td>45 (41–52)</td>
<td>42 (34.5–47.5)</td>
<td>35.5 (31.5–42)</td>
</tr>
<tr>
<td>QUEST</td>
<td>9 (9–12)</td>
<td>10 (9–12)</td>
<td>9 (9–13)</td>
<td>9 (9–12)</td>
</tr>
<tr>
<td>JAMA</td>
<td>0 (0–1)</td>
<td>0 (0–1.75)</td>
<td>1 (0–1)</td>
<td>0 (0–1)</td>
</tr>
</tbody>
</table>

3.2. Correlations

All three quality assessments correlated positively to each other, see Table 2. DISCERN subsection scores also correlated positively to each other. All readability scores correlated to each other positively, except for FRE which had a negative correlation, see Table 3.

Table 2. Kendall tau correlations between three quality assessment scores.

<table>
<thead>
<tr>
<th></th>
<th>DISCERN</th>
<th>QUEST</th>
<th>JAMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCERN</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QUEST</td>
<td>$\tau = 0.436 \ p &lt; 0.001$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>JAMA</td>
<td>$\tau = 0.3 \ p &lt; 0.001$</td>
<td>$\tau = 0.646 \ p &lt; 0.001$</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3. Kendall tau correlation between readability scores.

<table>
<thead>
<tr>
<th></th>
<th>GF</th>
<th>FKG</th>
<th>SMOG</th>
<th>FRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GF</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FKG</td>
<td>$\tau = 0.78 \ p &lt; 0.001$</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMOG</td>
<td>$\tau = 0.82 \ p &lt; 0.001$</td>
<td>$\tau = 0.85 \ p &lt; 0.001$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>FRE</td>
<td>$\tau = -0.61 \ p &lt; 0.001$</td>
<td>$\tau = -0.71 \ p &lt; 0.001$</td>
<td>$\tau = -0.625 \ p &lt; 0.001$</td>
<td>1</td>
</tr>
</tbody>
</table>

3.3. Kruskal-Wallis (KW) Test

Only DISCERN showed a significant difference in the KW test. This was also true for all subsection scores. Neither JAMA nor QUEST showed any difference. Based on the post hoc Games–Howell (GH) test, significant differences were identified (i) between bottom-ranked and top-ranked webpages with a mean difference of 9.968 (95% CI 5.7–14.2, $p < 0.001$) and (ii) between bottom-ranked and middle-ranked webpages with a mean difference of 5.5 (95% CI 0.21–10.8, $p < 0.04$). A GH test also identified significant differences within the subsection scores, which can be seen in Figure 1 and S2 data.

Regarding the subgroup analysis on sponsorship, no significant differences were identified for DISCERN and QUEST scores. However, for JAMA scores, there was a significant difference in the KW test (df = 6, $p = 0.005$). The GH test also showed a significant difference between health organizations and urologists with a mean difference of 1.042 (95% CI 0.006–2.077, $p = 0.048$).
with increasing inaccuracy or misinformation found online [22,23]. Yet, the internet remains a vital tool in a patient’s healthcare decision-making process [10].

There are only a handful of studies that have investigated the quality and readability of online health information, and the overall quality of online content related to urological treatment is considered poor [14,24]. None of the three quality assessment scores (IQR) based on median scores (IQR) were significant predictors for the DISCERN score. In the model, $R^2 = 0.215$, adjusted $R^2 = 0.151$, $F = 3.349$, $p = 0.002$. Top-ranked webpages were stronger predictors, $\beta = 10.515$ (95% CI 6.29–14.73), $p < 0.001$, compared to middle-rank webpages, $\beta = 5.823$ (95% CI 1.29–10.35), $p = 0.012$.

However, in contrast to the DISCERN score, a higher ranked webpage was a weaker predictor for both QUEST and JAMA score based on the regression model. In the QUEST regression model, $R^2 = 0.176$, adjusted $R^2 = 0.109$, $F = 2.61$, $p = 0.012$, where $\beta = 2.481$ (95% CI 0.4–4.56), $p = 0.02$.

In the JAMA regression model, $R^2 = 0.231$, adjusted $R^2 = 0.168$, $F = 3.681$ $p < 0.001$, where $\beta = 0.476$ (95% CI 0.01–0.944), $p = 0.046$. Notably, in this regression model, we also identified sponsorship from a urologist as a weak negative predictor for the JAMA score, where $\beta = -1.018$ (95% CI −1.72 to −0.312), $p = 0.005$. This latter finding was surprising.

With respect to readability, there was no significance identified in the regression analysis.

### 4. Discussion

An increasing number of patients are now relying on online resources to access and gather healthcare information for their medical needs. However, the overall quality of online health information is known to be poor [20,21]. It has become even more concerning with increasing inaccuracy or misinformation found online [22,23]. Yet, the internet remains a vital tool in a patient’s healthcare decision-making process [10]. There are only a handful of studies that have investigated the quality and readability of online health information regarding the surgical management of BPH [14,24]. This is the first study to assess both the quality and readability of online health information on BPH management, with a specific focus on HoLEP. Consistent with the existing literature, the overall quality of online content related to urological treatment is considered poor [14,24]. None of the three quality measures used in our study scored >60% in their total score.

In this study, the rank of the webpage is a significant positive predictor to the quality of health information. Linear regression analysis demonstrates that higher ranked webpages are more likely to score higher on all three quality assessment tools. Though in the KW test, ranking only had an influence on DISCERN, both in total and subsection scores. This is relevant to members of the general public as the literature suggests that patients often spend the majority of their time in top-ranked internet searches [25].
DISCERN is the suggested assessment tool in the health literature mainly because the questionnaire focuses on the reliability and details of the treatment [21,26]. In comparison, JAMA and QUEST tools focus on authorship and attribution [16,17]. This study highlights that many webpages lack the scientific literature or reliable authorship to support their claims; however, the literature has also shown that members of the general public do not spend a great deal of time identifying the attribution or reference of the webpage [25]. This can be alarming because members of the general public may perceive online health information to be of high quality; yet, the claims made in the online health information are not necessarily evidence-based practice.

The type of sponsorship has no influence on the quality of information. However, this linear regression analysis demonstrates that sponsorship from a urologist was a weak negative predictor of the JAMA score. This was a surprising finding, but it can be explained by the fact that the majority of the webpages lack good attribution and authorship and urologist as a category for sponsorship is the most common group.

The overall readability is moderately difficult; a minimum reading level of USA school grade 11 is needed. In Australia, there is no specific national target level for readability but some state health departments recommend patient education material should aim at a reading level of grade 6–8 [27]. This finding is consistent with previous studies where online information is written at a level that is above the average reader [27–29].

There are limitations to this study that should be addressed and allowed for future directions. Firstly, this study only focused on the quality of webpages which are presented in English and are written information available online. Webpages including videos and illustrations were not assessed in this study. Secondly, although 150 webpages were reviewed in this study, only 107 webpages (71%) were qualified for analysis. A larger sample size would potentially provide a more thorough examination and a more even distribution among top- to bottom-ranked webpages. Thirdly, the search term used was “enlarged prostate”, which was arbitrarily chosen as the selected layman search term. There is a possibility that using other relevant search terms such as “BPH” or “prostate hypertrophy” could have yielded a different set of webpages. Lastly, the Google search could have been limited by geographic location, as the initial search was conducted under an Australian IP address. The Google search may have favored or prioritized webpages within the Australian server. For future studies, it would be worthwhile conducting a study using IP addresses from different geographic locations, and different search terms, as well as including quality assessment of non-written information.

5. Conclusions

In conclusion, this study showed that online health information related to HoLEP and BPH is considered poor. Many webpages lack good evidence of their authorship and attributions. Sponsorship does not appear to have an influence on the quality of information. The overall readability can be considered as moderately difficult, such that a minimum reading level of USA school grade 11 is needed to comprehend online health information. Furthermore, this study demonstrated that a higher ranking of a webpage is a positive predictor for quality assessment tools such as DISCERN/QUEST/JAMA. Investment in improving the quality and readability of online health information is paramount, as there is an increasing reliance on online health information in a patient’s informed decision-making process.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/siuj5030029/s1, Table S1: MBS item statistics report and S2 data.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflict of interest.

References


25. Eysenbach, G.; Köhler, 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, 573–577. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.