Original Article

The Most Popular YouTube Videos About Shoulder Replacement Are of Poor Quality for Patient Education

Victor H. Martinez, B.S., Desiree Ojo, M.P.H., M.P.A., Jose M. Gutierrez-Naranjo, M.D., Mike Proffitt, Ph.D., and Robert U. Hartzler, M.D., M.S.

Purpose: To characterize the quality of YouTube total shoulder arthroplasty videos as a source of patient information using the DISCERN instrument. **Methods:** An analysis of the YouTube video library was performed, using a string of 6 search terms related to "total shoulder replacement" and "total shoulder arthroplasty" in the YouTube search engine. The first 20 videos from each search (n = 120) were selected. The top 25 most viewed videos were compiled, screened, and evaluated with the DISCERN score in the final analysis. Pearson's correlation coefficients were used to assess the correlation of DISCERN scores and video characteristics. Inter-rater reliability was calculated with the conger kappa score for multiple raters. **Results:** Twenty-five videos met inclusion criteria, 13 (52%) were produced by academic institutions, 7 (28%) by physicians, and 5 (20%) by commercial entities. Median total DISCERN score was 33 out of 80 (IQR: 28-44). The overall total DISCERN scores, showed no correlation with video likes or views and was negatively correlated with video power index ($r = -0.75$, $P = .001$). No association between total shoulder arthroscopy video source and DISCERN score could be demonstrated. All videos analyzed scored poorly by the DISCERN instrument. **Conclusions:** The current most popular shoulder replacement videos on YouTube are low-quality patient education resources. Furthermore, our study found no correlation between video popularity, as measured by the number of views and the DISCERN score. **Clinical Relevance:** Successful outcomes following total shoulder arthroplasty may be influenced by the quality of information patients receive.

Introduction

Modern-day medicine faces many challenges in providing adequate, understandable, and high-quality online patient education. This is mostly due to the increased time and effort required to regulate online educational resources, the variability of content, and low health literacy in the United States.[1,2] Furthermore, many healthcare providers may lack the necessary expertise in developing and providing adequate online educational resources.[3]

With more than 80% of patients obtaining health-related information from publicly available online resources, the Internet has become the dominant source for patients looking to learn about their medical conditions.[4] With patient education becoming more accessible on an individual level, understanding the material is vital for the patient to fully comprehend the necessary information to make informed decisions.[5] Patients and physicians should familiarize themselves with the quality of the resources due to the variety of

options available. High-quality preoperative patient education has been shown to reduce patient anxiety, improve pain management, and associated expectations, prevent harmful misunderstandings, and identify barriers to treatment that may be addressed before surgery.[6,7]

YouTube is the second-largest internet search engine behind Google and is one of the most popular and accessible informational resources on the Internet. However, because of YouTube's free and unregulated structure, the quality and content of publicly available health-related YouTube videos can be questionable.[7] Current orthopaedic studies on patient education YouTube videos have shown the content to be of poor quality.[7-15] Despite YouTube's popularity as a source of patient education, there is minimal information on the quality of online video content for total shoulder replacement.

The purpose of this study was to characterize the quality of YouTube total shoulder arthroplasty videos as a source of patient information using the DISCERN instrument. We hypothesized that the most popular videos on YouTube would not be of the highest quality and, therefore, lack essential information.

Methods

Search Method

An analysis of the YouTube video library was performed on June 14, 2022, using a string of 6 search terms related to "total shoulder replacement" and "total shoulder arthroplasty" (Fig 1) in the YouTube search engine. The first 20 videos from each search (n = 120) were selected. The search was conducted on the Chrome Browser using Incognito mode for cache clearance to minimize biased results for individual reviewers.

Inclusion and Exclusion Criteria

Videos shorter than 60 seconds, longer than 20 minutes, audio-only, not about total shoulder arthroscopy (TSA), and in languages other than English were excluded from this study.

Source of Videos

On the basis of their source, the videos were categorized as follows: 1) physician or physician group, 2) healthcare organization/hospital, and 3) commercial entity. Physician videos were uploaded by physicians or groups of physicians in private practices. Video uploads associated with medical centers, hospitals, and professional medical societies were defined as healthcare organization/hospital videos. Commercial videos were defined as those uploaded by a company with the intention of promoting and/or selling a product.

Video Characteristics

The following descriptive data were analyzed and documented for each video: 1) video length, 2) number of views, 3) number of days since upload, 4) view ratio (defined as views/day), 5) number of likes, 6) number of dislikes, 7) like ratio, (the number of likes × 100)/(the number of likes + number of dislikes), and 8) video power index. VPI is an index used to assess a video's popularity based on the number of views and likes, and it has been investigated by prior research.[8,15-18] The VPI is calculated by dividing (like ratio × view ratio) by 100. Since YouTube does not have a system for rating video popularity, the VPI has been used as an alternative.[18]

Video Quality and Content Analysis

A fellowship-trained orthopaedic shoulder surgeon and an orthopaedic surgery research resident used the DISCERN instrument to evaluate video quality and analysis. The DISCERN score is a validated tool that evaluates the quality of consumer information using a 16-question survey to provide a standard approach for evaluating the quality of written consumer health information on treatment options.[19,20] The 16 questions were categorized in 3 sections, with questions that evaluated 1) credibility, 2) quality and content, and 3) overall general score. These sections were scored from 1 (poor reliability) to 5 (highest reliability) for a total possible score of 80 points, yielding grades of "very poor" (16 to 28), "poor" (29 to 41), "fair" (42 to 54), "good" (55 to 67), and "excellent" (68 to 80). The reviewers (R.H. and J.G.) were blinded to the number of views per video, as they evaluated them in a randomized order. The reviewers then scored the videos independently, and the scores were averaged.

Statistical Analysis

Statistical analyses were performed in R version 4.2.0 using the RStudio integrated development environment version 2022.07.2 (Build 576).[21,22] Continuous variables are reported as medians and interquartile ranges, whereas categorical and ordinal variables are reported as frequencies and percentages of the total cohort. The corrr package version 0.4.4 was used to calculate Pearson's correlation coefficients among DISCERN scores and video characteristics.[23] Inter-rater reliability estimates for the DISCERN scores, and subscales were calculated using Conger's quadratic-weighted kappa for 2 raters using the irrCAC package version 1.0.[24]

Results

Overall Video Characteristics

One-hundred twenty videos were screened for inclusion and 45 met initial inclusion criteria. Secondary
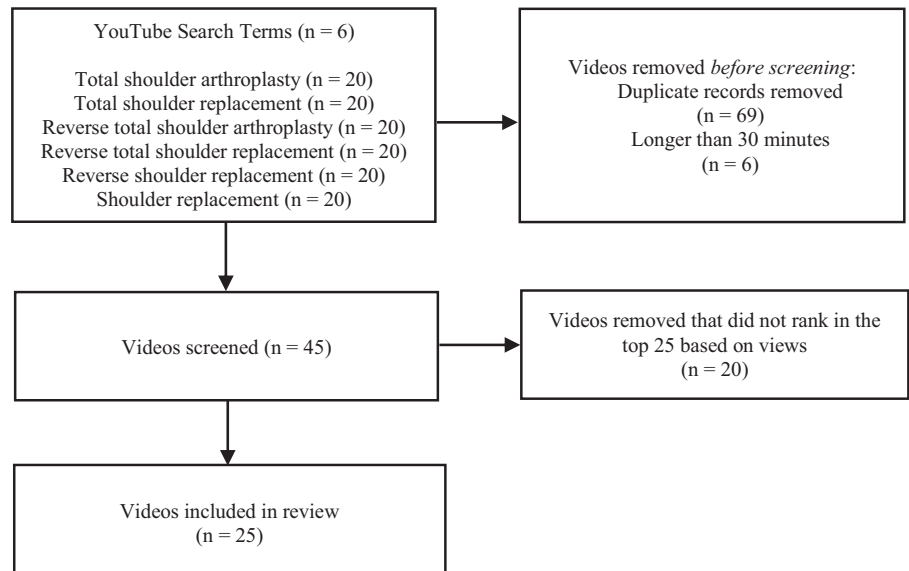
Fig 1. Flowchart of YouTube screening process.

to the constraint of author time and availability, the top 25 most viewed videos out of the pool of 45 were used. Thirteen of the top 25 videos were categorized as academic, 7 as physician, and 5 as commercial. The median number of views per TSA video was 34,891, with a high of 76,906 and a low of 14,791 (Table 1). The medium length of a video was 317 seconds (IQR: 225-641). The days since video upload was 1,693 (IQR: 1,245-2,693). Table 1 describes the remaining video properties.

Video Characteristics by Source

Commercial sources (n = 5) reported the highest median number of views per video with 97,156 (IQR: 29,404-237,515), whereas physician-produced videos (n = 7) had reported the lowest number of views, at 9,605 (IQR: 6,854-45,280) (Table 2). Physician videos had 2,872 days (IQR: 1,832-3084) on YouTube after video posting, the most of any source, while academic sources had the fewest at 1,513 (IQR: 1,090-1,693). Commercial videos had the longest median video

Table 1. TSA YouTube Video Descriptive Statistics

Video Characteristic	N = 25
Content Creator Source	
Academic	13 (52%)
Commercial	5 (20%)
Physician	7 (28%)
Video views	34,891 (14,791, 76,906)
Days since video posted	1,693 (1,245, 2,693)
View ratio	21 (13, 46)
Video likes	139 (84, 380)
Video dislikes	12 (5, 25)
Like ratio	95.1 (92.9, 97.1)
VPI like ratio	19 (12, 43)
Video length (seconds)	317 (225, 641)

Median n (%), (IQR).

length at 543 seconds (IQR: 225-641), followed by academic videos with a duration of 415 seconds (IQR: 231-840), while physician videos had the shortest video length at 242 seconds (IQR: 190-258). Table 2 includes additional information on videos by sources.

DISCERN Scores

The overall median DISCERN score for YouTube TSA videos was 33 (IQR: 28-44) of 80 possible points, as shown in Table 3. No videos scored a 5, which is the maximum points within the individual general score of the videos, while only 2 videos scored a 4/5. The overall total DISCERN scores, showed no correlation with video likes, or views and was negatively correlated with VPI ($r = -0.75$; $P = .001$). In fact, the DISCERN general score was negatively correlated with VPI, videos likes or views ($r = -.012$; $P = .001$). When categorized by video sources, the mean total DISCERN scores for YouTube TSA videos were low scoring and showed no significant differences between sources: physician 34 (29, 42), academic 33 (30, 45), and commercial videos 32 (25, 32) (Table 2).

Inter-rater Reliability

DISCERN scoring demonstrated substantial agreement with inter-rater reliability of 0.67 and 0.73 for the overall and general DISCERN scores, respectively.

Discussion

In this study, we found that the most commonly viewed videos about shoulder arthroplasty on YouTube are of low-quality for patient education. Furthermore, there was no correlation between popularity measured by the number of views and the DISCERN score. With more patients turning to the Internet for health

Table 2. Video Statistics and DISCERN by Sources

Video Characteristics	Academic, N = 13	Commercial, N = 5	Physician, N = 7
Video views	37,875 (27,305, 65,300)	97,156 (29,404, 237,515)	9,608 (6,854, 45,280)
Days since video posted	1,513 (1,090, 1,693)	1,944 (1,927, 2,328)	2,872 (1,832, 3,084)
View ratio	23 (18, 46)	50 (13, 80)	12 (4, 17)
Video likes	153 (126, 386)	2,000 (1,019, 7,000)	102 (65, 232)
Video dislikes	18 (8, 26)	72 (36, 390)	7 (2, 12)
Like ratio	94.7 (91.4, 96.4)	96.5 (95.5, 98.3)	95.7 (92.7, 97.5)
VPI like ratio	24 (16, 45)	77 (41, 456)	11 (4, 17)
Video length (seconds)	415 (231, 840)	543 (225, 641)	242 (190, 258)
Total DISCERN Score	33 (30, 45)	32 (25, 32)	34 (29, 42)
DISCERN Credibility Subscore	18.5 (18.0, 21.5)	18.5 (15.0, 20.5)	18.5 (15.5, 21.5)
DISCERN Quality Subscore	16.5 (12.0, 18.0)	10.5 (10.0, 12.0)	14.0 (12.8, 17.5)
DISCERN General Score			
1	3 (23%)	1 (20%)	1 (14%)
1.5	2 (15%)	2 (40%)	1 (14%)
2	1 (7.7%)	2 (40%)	2 (29%)
2.5	2 (15%)	0 (0%)	1 (14%)
3	1 (7.7%)	0 (0%)	1 (14%)
3.5	2 (15%)	0 (0%)	1 (14%)
4	2 (15%)	0 (0%)	0 (0%)

Median (IQR); n (%).

information and the prevalence of TSA surgeries rising at a faster rate than the most common orthopaedic procedures, such as total knee and total hip arthroplasty[25,26] high-quality, comprehensible online patient resources that adhere to a standard of accuracy are more important than ever.[4,6]

Previous research has highlighted the low quality and consistency of information accessible when examining educational health material on YouTube. Celik et al. examined the educational content on YouTube regarding rotator cuff repair (RCR).[27] Their evaluation of 67 videos using DISCERN found consistently poor scores. The average DISCERN score was 30.5, indicating that rotator cuff repair videos had essential and potentially serious shortcomings. In contrast, Ng et al. found that YouTube videos on total knee and hip arthroplasty videos scored an average of 54.7 on the DISCERN, deeming them of "fair-good" quality.[7] In 2025, 1,300,000 primary TKA and 650,000 primary THA procedures are projected to be performed in the United States versus 175,000-350,000 predicted TSA.[25,28] Thus, the differences in quality between TKA/THA versus TSA could be potentially attributed to more video producers for the former with an incentive to create high-quality content. It is evident that there is a need for improved TSA educational resources.[18]

An assessment of the quality of information regarding the return to sports following hip arthroscopy showed that physician uploads scored the highest in quality of content.[18] Additionally, Celik et al. showed that RCT videos by physicians had significantly higher DISCERN ($P < .001$) when compared with all other author types.[27] The overall quality of these videos demonstrates the potential importance and areas of improvement for video creators. Our investigation did not find a significant difference in the DISCERN score among video sources. Although, educational videos produced by physicians comprised 28% of all the TSA videos, they were not superior to the other sources. Interestingly, our study reported a video by Stanford Health Care ranked first with a 55 out of 80 (69%) score. Orthopaedic shoulder surgeons, either individually or through societies or other group efforts, should develop improved patient-focused videos published on YouTube that adhere to a standard of accuracy and quality to address many of the deficiencies identified in this study.[29]

Prior research has attempted to assess YouTube videos based on their VPI. In a study examining the quality of 122 YouTube videos on arteriovenous malformations (AVM), the videos with the greatest VPI had the lowest DISCERN score, indicating that the most popular videos on AVMs are not of the highest educational value.[30] Additionally, an evaluation of gastrectomy sleeve

Table 3. Overall TSA YouTube DISCERN Scores

DISCERN Score	33 (28, 44)
DISCERN Credibility Subscore	18.5 (16.5, 21.5)
DISCERN Quality Subscore	13.5 (12.0, 18.0)
DISCERN General Score	
1	5 (20%)
1.5	5 (20%)
2	5 (20%)
2.5	3 (12%)
3	2 (8.0%)
3.5	3 (12%)
4	2 (8.0%)

Median (IQR); n (%).

videos discovered that those produced by surgeons had a higher DISCERN score and a significantly lower VPI.[31] These findings support our study, as we discovered a negative correlation between VPI and total DISCERN score, and no correlation between the number of views and DISCERN score. The physician sources in our study had the lowest VPI score of the sources (11) but produced the highest quality of educational content. In contrast, commercial videos reported the highest VPI but the lowest DISCERN score. These findings indicate that VPI may be more suitable for entertainment than as a meaningful instrument for medical education.[30]

Limitations

This study has a few limitations. First, videos are frequently developed with various objectives, such as encouraging symptomatic individuals to seek medical care or providing insight on preoperative, intra-operative, and postoperative measures. Thus, the quality of the content may vary based on the target audience, making it a challenge to compare the overall quality of videos.[4] The study was limited by only analyzing the top 25 most viewed videos on YouTube. However, it has been reported that Internet users use only the top 10 search results.[32,33] Furthermore, the quality of the content of other orthopaedic-related videos is consistent with our findings. Lastly, some researchers propose establishing a more reliable methodology for evaluating YouTube video content quality.[34]

Conclusion

The current most popular shoulder replacement videos on YouTube are low-quality, patient education resources. Furthermore, our study found no correlation between video popularity, as measured by the number of views and the DISCERN score.

References

1. Rooney MK, Santiago G, Perni S, et al. Readability of patient education materials from high-impact medical journals: A 20-year analysis. *J Patient Exp* 2021:8. doi:10.1177/2374373521998847.
2. Fereidouni Z, Sabet Sarvestani R, Hariri G, Kuhpaye SA, Amirkhani M, Kalyani MN. Moving into action: The master key to patient education. *J Nurs Res* 2019;27:1-8.
3. Freda MC. Issues in patient education. *J Midwifery Womens Health* 2004;49:203-209.
4. Treffalls JA, Treffalls RN, Zachary H, et al. Quality analysis of online resources for patients with peripheral artery disease. *Ann Vasc Surg* 2022;83:1-9.
5. Wu JR, Holmes GM, DeWalt DA, et al. Low literacy is associated with increased risk of hospitalization and death among individuals with heart failure. *J Gen Intern Med* 2013;28:1174-1180.
6. Goyal R, Mercado AE, Ring D, Crijns TJ. Most YouTube videos about carpal tunnel syndrome have the potential to reinforce misconceptions. *Clin Orthop Rel Res* 2021;479:2296-2302.
7. Ng MK, Emara AK, Molloy RM, Krebs VE, Mont M, Piuzzi NS. YouTube as a source of patient information for total knee/hip arthroplasty: Quantitative analysis of video reliability, quality, and content. *J Am Acad Orthop Surg* 2021;29:e1034-e1044.
8. Erdem MN, Karaca S. Evaluating the accuracy and quality of the information in kyphosis videos shared on YouTube. *Spine (Phila Pa 1976)* 2018;43:E1334-E1339.
9. Cassidy JT, Fitzgerald E, Cassidy ES, et al. YouTube provides poor information regarding anterior cruciate ligament injury and reconstruction. *Knee Surg Sport Traumatol Arthrosc* 2018;26:840-845.
10. Macleod MG, Hoppe DJ, Simunovic N, Bhandari M, Philippon MJ, Ayeni OR. YouTube as an information source for femoroacetabular impingement: A systematic review of video content. *Arthroscopy* 2015;31:136-142.
11. Koller U, Waldstein W, Schatz KD, Windhager R. YouTube provides irrelevant information for the diagnosis and treatment of hip arthritis. *Int Orthop* 2016;40:1995-2002.
12. Addar A, Marwan Y, Algarni N, Berry G. Assessment of "YouTube" content for distal radius fracture immobilization. *J Surg Educ* 2017;74:799-804.
13. Staunton PF, Baker JF, Green J, Devitt A. Online curves: A quality analysis of scoliosis videos on YouTube. *Spine (Phila Pa 1976)* 2015;40:1857-1861.
14. Brooks FM, Lawrence H, Jones A, McCarthy MJH. Youtube™ as a source of patient information for lumbar discectomy. *Ann R Coll Surg Engl* 2014;96:144-146.
15. Kunze KN, Cohn MR, Wakefield C, et al. YouTube as a source of information about the posterior cruciate ligament: A content-quality and reliability analysis. *Arthrosc Sport Med Rehabil* 2019;1:e109-e114.
16. Springer B, Bechler U, Koller U, Windhager R, Waldstein W. Online videos provide poor information quality, reliability, and accuracy regarding rehabilitation and return to sport after anterior cruciate ligament reconstruction. *Arthroscopy* 2020;36:3037-3047.
17. Kunze KN, Alter KH, Cohn MR, et al. YouTube videos provide low-quality educational content about rotator cuff disease. *Clin Shoulder Elb* 2022;25:217-223.
18. Jildeh TR, Abbas MJ, Abbas L, Washington KJ, Okoroha KR. YouTube is a poor-quality source for patient information on rehabilitation and return to sports after hip arthroscopy. *Arthrosc Sports Med Rehabil* 2021;3:e1055-e1063.
19. Zade RT, Tartaglione JP, Chisena E, Adams CT, DiCaprio MR. The quality of online orthopaedic oncology information. *J Am Acad Orthop Surg Glob Res Rev* 2020;4:e19.00181.
20. The discern instrument. DISCERN. http://www.discern.org.uk/discern_instrument.php. Accessed September 17, 2022.
21. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed November 10, 2022.

22. RStudio Team. *RStudio*. Boston, MA: Integrated Development Environment for R. RStudio, PBC, 2022. version 2022.7.2.576). http://www.rstudio.com/. Accessed November 10, 2022.

23. Kuhn M, Jackson S, Cimentada J. *corrr: Correlations in R. R package* 2022. version 0.4.4. https://CRAN.R-project.org/package=corrr.

24. Gwet KL. *irrCAC: Computing chance-corrected agreement coefficients (CAC). R package version 1.0* 2019. https://CRAN.R-project.org/package=irrCAC.

25. Wagner ER, Farley KX, Higgins I, Wilson JM, Daly CA, Gottschalk MB. The incidence of shoulder arthroplasty: Rise and future projections compared with hip and knee arthroplasty. *J Shoulder Elbow Surg* 2020;29:2601-2609.

26. Farley KX, Wilson JM, Kumar A, et al. Prevalence of shoulder arthroplasty in the United States and the increasing burden of revision shoulder arthroplasty. *JB JS Open Access* 2021;6:e20.00156.

27. Celik H, Polat O, Ozcan C, Camur S, Kilinc BE, Uzun M. Assessment of the quality and reliability of the information on rotator cuff repair on YouTube. *Orthop Traumatol Surg Res* 2020;106:31-34.

28. Singh JA, Yu S, Chen L, Cleveland JD. Rates of total joint replacement in the United States: Future projections to 2020-2040 using the national inpatient sample. *J Rheumatol* 2019;46:1134-1140.

29. Schairer WW, Kahlenberg CA, Sculco PK, Nwachukwu BU. What is the quality of online resources about pain control after total knee arthroplasty? *J Arthroplasty* 2017;32:3616-3620.

30. Krakowiak M, Szmuda T, Fercho J, Ali S, Maliszewska Z, Słoniewski P. YouTube as a source of information for arteriovenous malformations: A content-quality and optimization analysis. *Clin Neurol Neurosurg* 2021;207:106723.

31. Ferhatoglu MF, Kartal A, Ekici U, et al. Evaluation of the reliability, utility, and quality of the information in sleeve gastrectomy videos shared on open access video sharing platform YouTube. *Obes Surg* 2019;29:1477-1484.

32. Goldenberg BT, Schairer WW, Dekker TJ, Lacheta L, Millett PJ. Online resources for rotator cuff repair: What are patients reading? *Arthrosc Sports Med Rehabil* 2019;1:e85-e92.

33. 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;324:573-577.

34. Azer SA. Are DISCERN and JAMA suitable instruments for assessing YouTube videos on thyroid cancer? Methodological concerns. *J Canc Educ* 2020;35:1267-1277.