

Evaluation of the Quality and Reliability of YouTube Videos on Gender Dysphoria

Ardıl Bayram Şahin¹, Oya Çelik²

¹ MD, Psychiatrist, Koç University, Graduate School of Health Sciences, Istanbul/Turkey

² Ms, Clinical Psychologist, Istanbul University, Istanbul Faculty of Medicine, Department of Psychiatry, Istanbul/Turkey

ABSTRACT

Purpose: Many people search for information on YouTube on almost every medical topic. The aim of this study was to assess the quality and reliability of YouTube videos concerning Gender Dysphoria.

Methods: We searched for Gender Dysphoria on YouTube on June 7, 2023 and conducted a detailed evaluation of the first 100 unique English videos by two expert observers. The videos were evaluated using modified DISCERN scale, Global Quality Scale (GQS), and Video Power Index (VPI).

Results: The kappa coefficient for interobserver agreement was 0.892. For all 100 videos, the median value of the modified DISCERN score was 2.5 (1-4.8), and GQS score was 3 (1-5), and 68% of the videos rated as poor to moderate reliability and quality (modified DISCERN score ≤ 3 , GQS ≤ 3). Reliable videos were uploaded by professional sources in a higher rate than unreliable videos (48% vs 12%, $p < 0.001$). Reliable videos had significantly higher modified DISCERN and GQS scores (3 vs 1.6, $p < 0.001$, 3.5 vs 2, $p < 0.001$; respectively). There were some significant correlations between some video characteristics and scores of quality and reliability scales. However, there was no significant difference between reliable and unreliable videos in terms of the popularity of video and no significant correlation between modified DISCERN and GQS scores and VPI ($p = 0.664$, $p = 0.201$, $p = 0.566$; respectively).

Conclusion: YouTube video quality for Gender Dysphoria was low to moderate, with a remarkable number of unreliable videos. There was no relation between video quality and popularity.

Keywords: Gender dysphoria, youtube, internet, video, quality, reliability

ÖZET

Amaç: Birçok kişi YouTube'da hemen hemen her tıbbi konuda bilgi aramaktadır. Bu çalışmanın amacı, Cinsiyet Hoşnutsuzluğu ile ilgili YouTube videolarının kalitesini ve güvenilirliğini değerlendirmektir.

Yöntemler: YouTube'da 7 Haziran 2023 tarihinde Cinsiyet Hoşnutsuzluğu ile ilgili arama yaptık ve ilk 100 İngilizce videoyu iki uzman gözlemci tarafından ayrıntılı olarak değerlendirdik. Videolar modifiye DISCERN, Global Kalite Ölçeği (GQS) ve Video Güç Endeksi (VPI) ile değerlendirildi.

Bulgular: Gözlemciler arası uyum için kappa katsayısı 0.892'dir. Yüz videonun tamamı için modifiye DISCERN skorunun medyan değeri 2,5 (1-4,8) ve GQS skoru 3 (1-5) olup videoların %68'i zayıf ila orta derecede güvenilir ve kaliteli olarak değerlendirilmiştir (modifiye DISCERN skoru ≤ 3 , GQS ≤ 3). Güvenilir videolar, güvenilir olmayan videolara kıyasla daha yüksek oranda profesyonel kaynaklar tarafından yüklenmiştir (%48'e karşı %12, $p < 0,001$). Güvenilir videoların modifiye DISCERN ve GQS puanları önemli ölçüde daha yüksektir (sırasıyla, 3'e karşı 1,6, $p < 0,001$, 3,5'e karşı 2, $p < 0,001$). Bazı video özellikleri ile kalite ve güvenilirlik ölççeklerinin puanları arasında bazı anlamlı korelasyonlar vardı. Ancak, güvenilir ve güvenilir olmayan videolar arasında videonun popülerliği açısından anlamlı bir fark bulunmamış ve modifiye DISCERN ve GQS puanları ile VPI arasında anlamlı bir korelasyon görülmemiştir (sırasıyla, $p = 0,664$, $p = 0,201$, $p = 0,566$).

Sonuç: Cinsiyet Hoşnutsuzluğu için YouTube video kalitesi düşük ila orta düzeydeydi ve kayda değer sayıda güvenilir olmayan video vardı. Video kalitesi ile popülerlik arasında bir ilişki bulunmamıştır.

Anahtar Kelimeler: Cinsiyet hoşnutsuzluğu, youtube, internet, video, kalite, güvenilirlik

Ardıl Bayram ŞAHİN
10000-0003-4515-7530

Oya ÇELİK
20000-0002-1832-2871

Correspondence: Ardıl Bayram ŞAHİN
MD, Psychiatrist, Koç University,
Graduate School of Health Sciences,
Istanbul/Turkey

Phone: +90 538 775 50 80

E-mail: ardilbayramsahin@gmail.com

Received: 20.11.2023

Accepted: 02.07.2024

Gender identity is a profound internal sense of affiliation with a man, or male, a woman, or female, or other gender (such as gender non-conforming, genderqueer, gender neutral, etc.) (1). The distress experienced by the individual due to the incongruence between their physical sex and gender identity is called gender dysphoria (2). The diagnosis of Gender Dysphoria (GD) was defined in the latest version of the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V) (3). Individuals with GD feels their gender identity that does not match the gender roles and societal expectations to their birth-assigned gender roles. This incongruence can lead to feelings of unease, restlessness, or internal conflict, which can have negative effects on mental health (4). The rate of people diagnosed with GD in the adult population is approximately one in a thousand individuals (3). Gender Dysphoria usually starts in early childhood, but it takes up to more than 20 years to seek transition (5,6). In order to alleviate the distress experienced by individuals due to GD, gender-affirming psychotherapy, hormone use, and surgical interventions may be recommended to appropriate individuals (7).

Individuals with GD and their families often turn to the internet for more information about GD (8). Video-sharing platforms also play a pivotal role in this information-seeking endeavor. YouTube is one of the most-watched internet platforms that has billions of users. Although it has significant potential to provide people with health-related information, the videos are uploaded not only by health professionals but also by non-professionals and there is no control mechanism for uploaded videos. There are numerous studies in the literature investigating the reliability and quality of YouTube videos addressing various health-related conditions (9,10,11,12). Previous research on YouTube videos showed that between one-third to half of videos about various medical conditions contain misleading information (13,14,9,15). Therefore, health-related videos should be questioned for accuracy, reliability, and quality before viewing.

However, there are currently no studies on the quality of videos on GD on YouTube of particular interest to individuals with gender dysphoria and their families. The aim of this study was to evaluate the content, reliability, and quality of YouTube videos on GD.

Material and Methods

Procedures

The YouTube website was searched for the term "gender dysphoria" using Google Chrome in incognito browser

mode without any filters (default mode) on June 7, 2023. The first 100 videos in English were included in the study. Two observers independently and simultaneously rated each video using the Qualtrics survey program. One of them was a psychiatrist and the other a clinical psychologist and both were EFS & ESSM certified psycho-sexologists. The videos were evaluated for reliability using the modified DISCERN scale, for quality using the Global Quality Scale (GQS), and for popularity using the Video Power Index (VPI). In addition, we assessed the degree of misinformation (containing information contradictory to established literature and/or current guidelines) in the videos using a 4-point scale (1-4), with numbers corresponding to none, low, medium, and high. This method has already been used in several previous studies (13,10,16). The YouTube videos were classified as reliable if they contained medically accurate information about GD. Medically accurate information about GD is seen as the definition, management, and social and medical transition processes of GD according to the DSM-V by the American Psychiatric Association, Standards of Care for the Health of Transgender and Gender diverse People by the World Professional Association for Transgender Health, and guidelines for psychological practice with transgender and gender nonconforming people by American Psychological Association (1,3,7). If a video contained both unreliable and reliable information, it was classified as unreliable.

Video characteristics such as the duration of the video, the total number of views, likes, dislikes, comments, and days, and the source of the video (uploader) and speaker of the video were recorded. Video Power Index that indicates the popularity of videos based on daily views, like, and dislike counts. The like rate was calculated using the formula $(\text{number of likes} \times 100) / (\text{number of likes} + \text{number of dislikes})$ and daily views were calculated using the formula $\text{total views} / \text{number of total days}$. Then, the Video Power Index of the video was calculated according to the formula $(\text{like rate} \times \text{daily views}) / 100$ based on the previous studies (10,11,17,18).

The uploaders of the videos were classified into professional sources (university, hospital, physician/psychologist) and non-professional sources (TV-YouTube channel, foundation, others). The speakers of the videos were classified as physician/psychologist, individuals with GD, others (YouTuber, TV speaker), and animation (no speaker).

Measurements

The DISCERN scale is a reliable and validated tool designed to help consumers of health information in assessing the quality of healthcare literature (19). It can be used by anyone without the requirement of specialized expertise. We used a modified 5-question DISCERN scale, with each question scored on a 5-point scale (16). The total score for the modified DISCERN is averaged and scaled to range between 1 and 5. Modified DISCERN scores within the range of video scores of 3-5 points indicate good, a score of 3 points indicate moderate, and scores 1-3 points indicate to poor reliability (Table 1). All videos were evaluated for the reliability of their content using the modified DISCERN scale.

Table 1: Modified DISCERN and Global Quality Scale		
Modified DISCERN		Reliability
1-Were aims clear and achieved?	1	Poor
2-Were the sources of information reliable?	2	Poor
3-Is the information balanced and unbiased?	3	Moderate
4-Are additional resources to learning provided?	4	Good
5-Does the video address areas of controversy/uncertainty?	5	Good
Global Quality Scale		Quality
Poor quality, very unlikely to be of any use to patients.	1	Poor
Poor quality but some information present, of very limited use to patients.	2	Limited
Suboptimal flow, some information covered but important topics missing, somewhat useful to patients.	3	Moderate
Good quality and flow, most important topics covered, useful to patients.	4	High
Excellent quality and flow, highly useful to patients	5	Excellent
<p><i>DISCERN scores: 1 and 2: poor reliability, 3: moderate reliability and 4 and 5: good reliability.</i></p> <p><i>Global Quality Scale: 1: poor quality, 2: limited quality, 3: moderate quality, 4: high quality and 5: excellent quality.</i></p>		

The GQS (Global Quality Score) scoring system which is introduced by Bernard et al. (2007) is used to assess the overall quality of the video (20). The scale provides information about the adequacy of the information in the video, the quality of that information, the overall flow of information, and how useful the author found the video for the audience. The observer should be qualified in the subject area. The quality of the video is evaluated through a 5-point scale. GQS scores within the range of 1-3 are classified as low quality, a score of 3 as medium quality, and a score of 4-5 as high quality (Table 1).

Statistics

Descriptive statistics of the data obtained from the study are given with median and range for numerical variables, while frequency and percentage analyses were employed for categorical variables. Modified DISCERN and GQS scores of the observers were averaged to calculate the mean scores. To assess the normal distribution of GQS and modified DISCERN scores, Kolmogorov–Smirnov and Shapiro–Wilk tests were conducted. Inter-rater agreement was determined using the kappa coefficient. Inter-rater reliability was measured by calculating the intraclass correlation coefficient. The data were not compatible with a normal distribution ($p < 0.05$). To explore the correlation between numerical variables, Spearman correlation analysis was implemented. In addition, for categorical variables featuring two groups, the Mann–Whitney U test was used to analyze GQS and modified DISCERN scores. The analyses were carried out with SPSS 22.0 software with a selected significance level of $p < 0.05$.

Ethical approval was not obtained since the study utilized publicly accessible videos and no human or animal subjects were used. Since all data used in this study were publicly available, permission from YouTube was not required. All confidential information, such as the name of the uploader, was maintained in strict confidence.

Results

The first 100 unique English videos about GD on YouTube were analyzed. One video was excluded due to duplication, while 5 videos were excluded owing to non-English content. The level of agreement between the 2 observers when classifying the videos as reliable and unreliable was highly positive (kappa coefficient= 0.892). The intraclass correlation coefficient was 0.908 for modified DISCERN and 0.873 for GQS, indicating a great reliability between the 2 observers for both tools.

Among the 100 videos, 61 videos were categorized as reliable, while 39 of them as unreliable that including inaccurate or misleading information concerning GD. The duration of the video, the number of followers, total views, likes, dislikes, and comments, daily views, like rate, VPI, source of video, and speaker of the video were shown in

Table 2. A notable rate of videos (67%) was uploaded by non-professional sources. Sixty-eight % of the videos rated as poor to moderate reliability and quality (modified DISCERN score ≤ 3 , GQS ≤ 3). For all 100 videos, the median value of the modified DISCERN score was 2.5(1-4.8), and GQS score was 3(1-5).

Table 2: Comparison of video engagement metrics and evaluation scales according to reliable and unreliable videos.

	Total (n=100) Median (range) or n (%)	Reliable (n=61) Median (range) or n (%)	Unreliable (n=39) Median (range) or n (%)	Z / χ^2	P
Characteristics of videos					
Duration (m)	10 (1-114)	9.7 (1-114)	11.3 (1.7-95.8)	-1.707	0.088
Number of followers	28,000 (36-38,400,000)	28,000 (36-38,400,000)	11,300 (36-38,400,000)	-1.313260	0.189
Number of total views	7,990 (100- 1741892)	13,815 (100-1741892)	5,243 (270-1330230)	-1.862150	0.063
Number of likes	342 (0-201,851)	537 (0-201,851)	181 (8-75,694)	-2.010605	0.044
Number of dislikes	13.5 (0-5137)	16 (0-3698)	7 (0-5137)	-1.217043	0.224
Number of comments	104.5 (0-13456)	141 (0-12521)	53 (1-13456)	-1.067717	0.286
Number of total days	737,5 (17-4897)	866 (17-4897)	591 (20-2506)	-0.872344	0.034
Number of daily views	144.443 (0233-18773,177)	18,440 (0,233-18773,177)	11.825 (0.295-1068,650)	-0.462887	0.643
Like rate	97.893 (0-100)	98.371 (0-100)	96.154 (50-100)	-0.998993	0.318
Video Power Index	13.741 (0-1540.171)	15.549 (0-15400.171)	10.706 (0.204-1021.875)	-0.434619	0.664

Uploader of video					
Professional account	33	29 (48%)	4 (12%)	14.957674	< .001
Non-professional account	67	32 (52%)	35 (88%)		
Speaker of video					
Professionals	52	29 (47.5%)	23 (59%)	8.417280	0.038
Subject	33	25 (41%)	8 (20.5%)		
Others	10	3 (5%)	7 (8%)		
Animation (no speaker)	5	4 (6.5%)	1 (2.5%)		

Scales					
Modified DISCERN	2.5 (1-4.8)	3.00 (1.8-4.8)	1.6 (1-3.6)	-6.429754	< 0.001
Global Quality Scale	3 (1-5)	3.5 (2-5)	2.0 (1-3)	-7.496546	< 0.001

Values of $p < 0.05$ were accepted as significant and marked in bold.

The number of likes and total days were significantly higher in the reliable videos than in the unreliable ($p < 0.05$). However, there was no significant difference between reliable and unreliable groups according to the VPI value ($p = 0.664$). Reliable videos were uploaded by professional sources in a higher rate than unreliable videos, and unreliable videos were uploaded by non-professional sources at a higher rate than reliable sources (48% vs 12%, $p < 0.001$, 88% vs 52%, $p < 0.001$; respectively). There were significant differences in speaker type of the video between reliable and unreliable groups ($p < 0.05$). Especially the proportions of physician/psychologist and individuals with GD as speakers in the reliable group differed from the proportions in the unreliable group (%47.5 and %41 vs %59 and

%20.5). Reliable videos had a significantly higher modified DISCERN and GQS scores than the unreliable videos (3 vs 1.6, $p < 0.001$, 3.5 vs 2, $p < 0.001$; respectively).

Videos uploaded by professional sources had greater modified DISCERN and GQS scores ($p < 0.001$, $p < 0.001$; respectively) in comparison to those uploaded by non-professional sources (Table 3). There were significant correlations between modified DISCERN score and number of followers, total views, likes, dislikes, comments, and total days. Similarly, there were correlations between the GQS score, and the number of total views and total days. However, there was no significant correlation between modified DISCERN and GOQ scores and the VPI value ($p = 0.201$, $p = 0.566$; respectively) (Table 4).

Table 3: Comparison of video evaluation tool scores according to professional and non-professional sources.

Scales	Professional n=33 Median (range)	Non-professional n=67 Median (range)	Z	P
Modified DISCERN	3.8 (1-4.8)	2.3 (1.1-4.2)	-5.921161	< 0.001
Global Quality Scale	3.5 (1-5)	2.5 (1-4)	-4.540107	< 0.001

Values of $p < 0.05$ were accepted as significant and marked in bold.

Table 4: Correlation analyses for Global Quality Scale and modified DISCERN scores of the videos.

Scales	Modified DISCERN		Global Quality Scale	
	r	p	r	p
Duration	-0.016	0.874	0.190	0.848
Number of followers	0.199597	0.046	0.093731	0.354
Number of total views	0.265951	0.007	0.234169	0.019
Number of likes	0.215394	0.031	0.192575	0.055
Number of dislikes	0.231448	0.020	0.168928	0.093
Number of comments	0.203002	0.049	0.129506	0.214
Number of total days	0.207169	0.038	0.232559	0.012
Number of daily views	0.134341	0.182	0.048907	0.629
Like rate	-0.121679	0.227	-0.064007	0.527
Video Power Index	0.129001	0.201	0.058037	0.566

Values of $p < 0.05$ were accepted as significant and marked in bold.

Discussion

In this study, we aimed to analyze the 100 most relevant videos concerning GD on the YouTube channel to evaluate these videos in terms of reliability and quality. Previous studies have evaluated the quality and reliability of YouTube videos for different medical conditions, such as premature ejaculation, erectile dysfunction, rheumatologic disease, body dysmorphic disorder, gender-affirming surgery, self-breast examination, and rheumatoid arthritis (9,10,11,13,14,17,18). To our knowledge, this was the first in-depth study to assess the reliability and quality of information regarding GD on YouTube.

Our results showed that 39% (n=39) of videos that containing misleading and/or unreliable information were identified as unreliable and, 68% (n=68) of videos exhibited a level of reliability and quality ranging from poor to moderate based on modified DISCERN and GQS scores. We found that reliable videos had significantly higher quality and reliability scores than unreliable videos. Consistent with our study, two separate studies that analyzed the quality of information on YouTube regarding premature ejaculation and erectile dysfunction, demonstrated that 30% and, 28% of the videos, respectively, were not reliable (10,13). In another study analyzing the quality and reliability of YouTube videos on a rheumatologic disease indicated that 46% of the videos were of low to moderate reliability and 56% of low to moderate quality (11). In a study that analyzed erectile dysfunction videos on TikTok, it was reported that 80% of the videos were not reliable (16). Another study of Vulvodynia videos on YouTube reported that 58% of the videos were of low quality (21). A recent systemic review assessing the reliability of health-related videos on YouTube showed that YouTube is not a reliable source of medical and health-related information (22). A study reported that many popular YouTube videos about prostate cancer contain biased and/or low-quality information (23). Currently, there were reported some misconceptions and scientific misinformation about gender dysphoria and gender-affirming care. (24,25).

We found that reliable videos were uploaded by professional and non-professional sources with nearly equal proportions, while unreliable videos were mostly uploaded by non-professional sources. Our study findings were in accordance with findings of previous studies that showed the better video quality in content uploaded by

als with GD coming second. Although the proportion of physician /psychologists as speakers in unreliable videos was higher than in reliable videos, the proportion of individuals with GD as speakers was higher in reliable videos. In previous studies, it was shown that video uploaders were primarily professional sources, and secondarily patients (10,22,27). A study that analyzed YouTube videos on gender affirmation surgery showed that videos uploaded by professional sources were of higher quality than compared to those uploaded by patients (17). Principally, professional sources should provide more reliable videos. It was found that videos uploaded by professional sources had greater reliability and quality than those uploaded by non-professional sources. Nevertheless, our investigation revealed that physician/psychologists also played a role in generating unreliable content. These results were similar to outcomes observed in a few studies evaluating the effects of YouTube videos (13,14). We found that unreliable videos have a higher rate of physician/psychologists as speakers, thereby the source of the uploader is more important for reliability.

When we compared the reliable and unreliable groups based on video characteristics, we found that the reliable videos had a higher number of likes and total days. However, there was no difference between the two groups in terms of video popularity. Additionally, we showed correlations between reliability scores and the number of views, likes, dislikes, comments, followers, and total days. Similarly, there were some correlations between quality scores and the number of views, and total days. However, there was no correlation between reliability and quality, and video popularity. Consistent with our results, previous studies have similarly not found a relationship between video popularity and quality (10,18,22). Furthermore, another study revealed an inverse relationship between video quality and popularity (23). These results showed that unreliable videos that disseminated misinformation were as popular as reliable videos.

YouTube, boasting an extensive user base, stands as one of the most prevalent social media platforms. With the widespread use of the internet, the number of people watching video content on these platforms is steadily increasing. Individuals dealing with gender dysphoria often face obstacles when trying to access crucial healthcare services (28). As a result, they might choose to navigate social media platforms in search of information about their dysphoria, aiming to avoid instances of discrimination.

Moreover, these platforms are used for sharing experiences where individuals can benefit from the experiences of other individuals with gender dysphoria (29). However, health-related information gleaned from the internet can be incomplete and misleading (9,22). Although it is impossible to control all the data available on YouTube, further research into assessing video quality could lead people to use these platforms more appropriately.

This study has some limitations. Firstly, we only included videos in English, therefore generalizations cannot be possible in other languages. Additionally, there is no clear consensus on how to evaluate the quality of the videos on health-related videos. However, we used two scales that are commonly used in previous studies (22). Finally, we should also acknowledge the possibility of obtaining different results when conducting similar searches at different times. Despite these limitations, the current study can be considered an assessment of the GD on YouTube videos at a specific time.

Conclusion

In this study, it was determined that YouTube videos concerning GD are mainly uploaded by non-professional sources and exhibiting a range of low to moderate video quality. More than one of third of videos are unreliable that contain misleading information. There is no relation between the video quality and popularity. The quality of the video is not always be determined by the speaker; the source of the video is very crucial for video quality. Therefore, universities, health organizations and other professional sources specialized in this field should produce high quality videos with accurate insights into GD.

Declarations

Disclosure of Interest

The authors declare that they have no conflict of interest.

Funding

No financial support or scholarship has been received for this research.

Ethics Committee Approval

This study was exempt from ethical approval of the study institution because it involved the use of public access data only.

Data-Sharing

The data that support the findings of this study are available from the corresponding author with request.

Authors' Contributions

Concept – A.B.Ş., O.Ç.; Design – A.B.Ş., O.Ç.; Supervision – A.B.Ş.; Resources – A.B.Ş., O.Ç.; Materials – A.B.Ş., O.Ç.; Data Collection and Processing - A.B.Ş., O.Ç.; Analysis and/or Interpretation A.B.Ş., O.Ç.; Literature Review – A.B.Ş., O.Ç.; Writing – A.B.Ş., O.Ç.; Critical Review - A.B.Ş., O.Ç.

References

1. American Psychological Association. Guidelines for psychological practice with transgender and gender nonconforming people. *Am Psychol.* 2015;70:832-864. <https://doi.org/10.1037/a0039906>
2. Fisk NM. Gender dysphoria syndrome (the how, what and why of a disease), Proceedings of the second interdisciplinary symposium on gender dysphoria syndrome. California, Stanford, 1973.
3. American Psychiatric Association. Diagnostic and statistical manual of Mental Disorders (5th ed.). Washington D.C: American Psychiatric Association. 2013. <https://doi.org/10.1176/appi.books.9780890425596>
4. Testa RJ, Habarth J, Peta J, et al. Development of the Gender Minority Stress and Resilience Measure. *Psychol Sex Orientat Gen Divers.* 2015;2:65–77. <https://doi.org/10.1037/sgd0000081>
5. Ruble DN, Taylor LJ, Cyphers L, et al. The role of gender constancy in early gender development. *Child Dev.* 2007;78:1121-1136. <https://doi.org/10.1111/j.1467-8624.2007.01056.x>
6. Zaliznyak M, Yuan N, Bresee C, et al. How Early in Life do Transgender Adults Begin to Experience Gender Dysphoria? Why This Matters for Patients, Providers, and for Our Healthcare System. *Sex Med.* 2021;9:100448. <https://doi.org/10.1016/j.esxm.2021.100448>
7. Coleman E, Radix AE, Bouman WP, et al. Standards of Care for the Health of Transgender and Gender Diverse People, Version 8. *Int J Transgend Health.* 2022;23(Suppl.1):S1-S259. <https://doi.org/10.1080/026895269.2022.2100644>
8. Lena Herrmann, Carola Bindt, Sarah Hohmann, et al. Social media use and experiences among transgender and gender diverse adolescents. *International Journal of Transgender Health.* Published online 04 Sep 2023. (<https://doi.org/10.1080/026895269.2023.2252410>)
9. Warren CJ, Wisener J, Ward B, et al. YouTube as a Patient Education Resource for Male Hypogonadism and Testosterone Therapy. *Sex Med.* 2021;9:100324. <https://doi.org/10.1016/j.esxm.2021.100324>
10. Fode M, Nolsøe AB, Jacobsen FM, et al. Quality of Information in YouTube Videos on Erectile Dysfunction. *Sex Med.* 2020;8:408-413. <https://doi.org/10.1016/j.esxm.2020.05.007>
11. Karakoyun A and Yildirim A. YouTube videos as a source of information concerning Behçet's disease: a reliability and quality analysis. *Rheumatol Int.* 2021;41:2117–2123. <https://doi.org/10.1007/s00296-021-05009-9>
12. Civan Kahve A, Asut G, Kaya H, et al. What does YouTube® say about schizophrenia: Is it a reliable source of information?. *Acibadem Univ Sağlık Bilim Derg.* 2022;13(1):67-73. <https://doi.org/10.31067/acusaglik.981253>
13. Esen E, Aslan M, Sonbahar BÇ, et al. YouTube English videos as a source of information on breast self-examination. *Breast Cancer Res Treat.* 2019;173:629-635. <https://doi.org/10.1007/s10549-018-5044-z>
14. Gul M and Diri MA. YouTube as a Source of Information About Premature Ejaculation Treatment. *J Sex Med.* 2019; 16:1734-1740. <https://doi.org/10.1016/j.jsxm.2019.08.008>
15. Leong AY, Sanghera R, Jhaji J, et al. Is YouTube useful as a source of health information for adults with type 2 diabetes? A South Asian perspective. *Can J Diabetes.* 2018;42:395-403.e4. <https://doi.org/10.1016/j.jcjd.2017.10.056>

16. Bernstein A, Zhu M, Loloi J, et al. TikTok as a source of information regarding premature ejaculation: a qualitative assessment. *Sex Med.* 2023;11(2):qfac020. <https://doi.org/10.1093/sexmed/qfac020>
17. Batur AF, Sun H, Pope R, et al. A comprehensive analysis of YouTube videos on gender affirmation surgery. *Int Urogynecol J.* 2023;34:2275-2283. (<https://doi.org/10.1007/s00192-023-05542-0>)
18. Dobosz M, Lewandowski M, Świerczewska Z, et al. Are YouTube videos a reliable source of information about body dysmorphic disorder?. *Postep Psychiatr Neurol.* 2023;32:76-82. <https://doi.org/10.5114/ppn.2023.128706>
19. Charnock D, Shepperd S, Needham G, et al. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health.* 1999;53:105-111. <https://doi.org/10.1136/jech.53.2.105>
20. Bernard A, Langille M, Hughes S, et al. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. *Am J Gastroenterol.* 2007;102:2070-2077. <https://doi.org/10.1111/j.1572-0241.2007.01325.x>
21. Aglamis SO, Senel S and Koudonas A. Quality analysis of YouTube videos on vulvodynia. *Sex Med.* 2023;11:qfac013. <https://doi.org/10.1093/sexmed/qfac013>
22. Osman W, Mohamed F, Elhassan M, et al. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med Educ.* 2022;22:382. <https://doi.org/10.1186/s12909-022-03446-z>
23. Loeb S, Sengupta S, Butaney M, et al. Dissemination of Misinformative and Biased Information about Prostate Cancer on YouTube. *Eur Urol.* 2019;75:564-567. <https://doi.org/10.1016/j.eururo.2018.10.056>
24. McNamara M, Lepore C, Alstott A, et al. Scientific Misinformation and Gender Affirming Care: Tools for Providers on the Front Lines. *J Adolesc Health.* 2022;71:251-253. <https://doi.org/10.1016/j.jadohealth.2022.06.008>
25. Lepore C, Alstott A, McNamara M. Scientific Misinformation Is Criminalizing the Standard of Care for Transgender Youth. *JAMA Pediatr.* 2022;176:965-966. <https://doi.org/10.1001/jamapediatrics.2022.2959>
26. Om A, Ijeoma B, Kebede S, et al. Analyzing the quality of aesthetic surgery procedure videos on TikTok. *Aesthet Surg J.* 2021;41:2078-2083. <https://doi.org/10.1093/asj/sjab291>.
27. Cetin A. Evaluation of YouTube Video Content Related to the Management of Hypoglycemia. *Cureus.* 2021;13:e12525. 2021 Jan 6. <https://doi.org/10.7759/cureus.12525>.
28. Kachen A and Pharr JR. Health Care Access and Utilization by Transgender Populations: A United States Transgender Survey Study. *Transgend Health.* 2020;5:141-148. <https://doi.org/10.1089/trgh.2020.0017>
29. McInroy LB and Craig SL. Transgender representation in offline and online media: LGBTQ youth perspectives. *J Hum Behav Soc Environ.* 2015;25:606-617. <https://doi.org/10.1080/10911359.2014.995392>