ARTICLE
Analysis of quality information provided by “Dr. YouTube™” on Phimosis

Simone Cilio1, Claudia Collà Ruvolo1, Carmine Turco1, Massimiliano Creta1, Marco Capece1, Roberto La Rocca1, Giuseppe Celentano1, Gianluigi Califano1, Simone Morra1, Alberto Melchionna1, Francesco Mangiapia1, Felice Crocetto1, Paolo Verze2, Alessandro Palmieri1, Ciro Imbimbo1 and Vincenzo Mirone1

© The Author(s), under exclusive licence to Springer Nature Limited 2022

The objective of the current study was to evaluate the quality of the information provided in YouTube™ videos on phimosis. The term “phimosis” was searched on YouTube™, and the Patient Education Materials Assessment Tool (PEMAT) for Audio/Visual Materials (Understandability and Actionability sections, good-quality score of minimum 70%) and misinformation scale (rated from 1 to 5) were used to assess video quality. Quality assessment was investigated over time. Of all, 60 were eligible for analysis. Healthcare providers were the authors of 75.0% of the videos, and 73.3% of the videos were patient-targeted. The median Understandability score was 42.9% (interquartile range [IQR]:34.5–58.9) and ranged from 28.6 to 42.9% (2013–2020). The median Actionability score was 50.0% (IQR:25.0–56.2) and ranged from 25.0 to 50.0% (2013–2020). The median misinformation score was 2.8/5 (IQR:1.6–3.6), and although the score fluctuated over time, the median score was 2.6 both in 2013 and in 2020. According to our results, although an increase of PEMAT over time was observed, the overall quality of the information uploaded on YouTube™ is low. Therefore, at present, YouTube™ cannot be recommended as a reliable source of information on phimosis. Video producers should upload higher-quality videos to help physicians and patients in the decision-making process.

IJIR: Your Sexual Medicine Journal (2023) 35:398–403; https://doi.org/10.1038/s41443-022-00557-5

INTRODUCTION
Phimosis is defined as an inability to pull the foreskin back behind the glans [1, 2]. This condition is considered pathological when caused by a distal fibrous ring around the prepuce [3].
Phimosis is diagnosed in 0.5–13.0% of men, regardless of age [4]. It is considered physiological in childhood. However, only 1% of phimotic foreskins remain non-retractile during the teenage years, becoming pathological [3]. The surgical procedure used to treat phimosis is circumcision [1]. In the general population, a reluctance to visit a urologist/andrologist regarding genital issues has been noted [5]. As a consequence, patients prefer to get information on their condition through the Internet [6]. Particularly, social media has become an important tool, demonstrating multiple advantages for the sharing and dissemination of scientific knowledge among the healthcare community [7]. This phenomenon has been even more emphasized during the SARS-CoV2 pandemic and has led to delays in outpatient care and surgery in urology [8, 9].

Today, among all the Internet sources, YouTube™ is asserted as the second most clicked site worldwide [10]. It is a social media platform used to share and obtain information with more than one billion registered users and almost five billion videos watched daily [11]. The largest age group of YouTube™ users is 25 years old [12].

Several previous studies on YouTube™ information on urological [13–15] and non-urological topics [16–18] have already been published, showing a potentially harmful role of YouTube™ videos on the general population, especially those without a medical background. For example, several articles have been published on the low quality information of YouTube™ videos regarding erectile dysfunction and premature ejaculation [13, 14].

To the best of our knowledge, no prior research has examined the quality of information uploaded on YouTube™ concerning phimosis. Thus, the current study aimed to evaluate the quality of information of YouTube™ videos on this condition, in order to make internet users aware of the potential role of YouTube™ videos as learning, as well as supporting tools for healthcare providers in the decision-making process.

MATERIALS AND METHODS
Search strategy and video selection criteria
On January 16th, 2021 from 4.50 p.m. to 8.30 p.m. UTC + 0, we used the term “phimosis” on the YouTube™ search engine. To avoid single user bias, “incognito status” was used, the log out from any personal accounts was performed, and a proxy located in London (United Kingdom), via a Virtual Private Network (VPN) software, was used. The default YouTube™ “setting search” does not apply any filter and sort the results by relevance. To provide the most reliable results, the same research was performed on Google Chrome, Microsoft Edge, Mozilla Firefox, and Opera Browsers and the same results were achieved.

The first 100 videos were selected by one doctor (a junior urology resident [T.C.]), assuming that more than 90% of Internet users watch the
videos within the first three pages (60 videos) [19]. The following exclusion criteria were applied: (i) videos not in the English language; (ii) videos not relating to penile phimosis; and (iii) videos longer than 60 min. In case of a duplicate, only one was considered, and if one video was a part of a compilation, it was taken into account as a single video.

For each video, the following variables were collected: length (seconds), number of views, persistence on YouTube™ (months), number of thumbs-up, number of thumbs-down, number of total comments, number of comments disabled, medical advice requested in comments, channel subscribers, video author (healthcare providers vs patients vs others [such as pedagogist or medical students]), video author's country (defined as India vs the rest of the world [Australia, Brazil, France, Greece, Italy, Malaysia, Pakistan, Ukraine, UK, and the USA]), video target (healthcare providers vs patients), video topic (medical lesson [focused on definition and diagnosis of phimosis, pathological implications, and at least one therapeutic option] vs patients experience vs information for patients vs commercials), and video upload year (2013–2017 vs 2018–2020).

Quality and misinformation assessment tools
The quality of videos was assessed by two investigators (a junior urology resident [S.C] and a senior urology resident [C.G.]). Any potential conflicts or queries were independently reviewed by a separate author (an Associate Professor [A.P.]). A single final score was assigned for each video.

The following quality assessment tools were used: Patient Education Materials Assessment Tool for Audio/Visual materials (PEMAT A/V) and a 5-item misinformation scale.

The PEMAT A/V is a systematic tool developed by the Agency for Healthcare Research and quality to evaluate and compare the Understandability and Actionability of patient education materials [20]. It consists of 13 items measuring Understandability and four items measuring Actionability. Each item is evaluated as Agree (1 point), Disagree (0 points), or Not applicable (N/A). The PEMAT A/V provides two separate scores for Understandability and Actionability. Each score is the percentage obtained by the sum of the total points, divided by the number of items evaluated as Agree or Disagree. The higher the score, the more understandable or actionable the material was [21]. In particular, for each score, the cut-off to define good video quality is 70% [21]. A 5-item misinformation scale [22] was specifically created for this study according to the European Association of Urology (EAU) Patient Information [23]. Specifically, misinformation items were: (1) Accurately explains the topic; (2) Factually describes when the condition is pathological; (3) Provides correct explanation of the surgical option; (4) Mentions the EAU approved recommendations; (5) Presents and explains treatment results clearly. Each item was rated from 1 to 5 (1-Strongly Disagree; 2-Disagree; 3-neither agree nor disagree; 4-Agree; 5-Strongly Agree). The results were presented as mean (Standard Deviation [SD]) and median (interquartile range [IQR]) and ranged from 1 to 5. The lower the score, the higher the misinformation level was.

Video quality assessment overtime
Locally Estimated Scatterplot Smoothing (LOESS) methodology was used to graphically explore median change over time (2013–2020) for PEMAT A/V Understandability, PEMAT A/V Actionability, and misinformation score.

Statistical analyses
Descriptive statistics are presented as means, medians, and IQR for continuously coded variables or counts and percentages for categorically coded variables. Chi-square test, t-test, and Kruskal–Wallis tests examined the statistical significance in proportion, mean and median differences. Cohen kappa statistics were used to measure the reliability of the investigator's evaluations of the videos. To examine potential correlations between variables, Pearson's correlation was used. In all statistical analyses, R software (www.r-project.org) environment for statistical computing and graphics (R version 4.0.0), and Microsoft Excel 2009 were used. All tests were two-sided with a level of significance set at p < 0.05.

RESULTS
Videographic characteristics
Of all 100 videos examined, 60 were selected for analysis (Fig. 1). The median length was 195 s (IQR:91.2–402.5; range:15–994), the median number of views was 17269 (IQR:2343.2–86700.0; range:87–1058974), and the median persistent on YouTube™ was 17 months (IQR:6.0–38.5; range:1–164). Moreover, across the sample, the median number of thumbs-up, thumbs-down, comments, and channel subscribers were 77.5 (IQR:9.8–297.0; range:1–2274), 10 (IQR:1.8–38.5; range:0–776), 6.5 (IQR:0.0–86.5; range:0–712), and 19950 (IQR:3370.0–50900.0; range:0–894000), respectively. In 8 videos (13.3%) comments were disabled, and in 30 videos (50.0%) medical advice requests in comments were recorded. Among all the videos, 45 (75.0%) were produced in India, while 15 (25.0%) were produced in the rest of the world (considered as a single group due to excessively small videos.
number in each country recorded). Of all the videos, 75.0% (n = 45), 20.0% (n = 12), and 5.0% (n = 3), were produced by healthcare providers, patients, or others, respectively. Additionally, 26.7% (n = 16) and 73.3% (n = 44) were targeted towards healthcare providers and patients, respectively. Medical lesson was the main topic addressed (n = 28,46.7%), followed by information for patients (n = 18,30.0%), patient experience (n = 12,20.0%), and commercials (n = 2,3.3%). Finally, 73.3% of videos were uploaded from 2018 to 2020 (Table 1).

**Table 1.** Videographic characteristics of 60 YouTube™ videos on “phimosis” selected on January 16th, 2021.

| Videographic characteristics | Value |
|------------------------------|-------|
| Length (sec)                 | Median (IQR) 195.0 (91.2–402.5) |
|                              | Range 15–994 |
| Number of views              | Median (IQR) 17269.0 (2343.2–86700.0) |
|                              | Range 87–1058974 |
| Persistence on YouTube™ (months) | Median (IQR) 17.0 (6.0–38.5) |
|                              | Range 1–164 |
| Thumbs-up, n                 | Median (IQR) 77.5 (9.8–297.0) |
|                              | Range 1–2274 |
| Thumbs-down, n               | Median (IQR) 10.0 (1.8–38.5) |
|                              | Range 0–776 |
| Comments, n                  | Median (IQR) 6.5 (0–86.5) |
|                              | Range 0–712 |
| Disabled comments, n (%)     | No 52 (86.7) |
|                              | Yes 8 (13.3) |
| Medical advice requested in comments, n (%) | N/A 14 (23.3) |
|                              | No 16 (26.7) |
|                              | Yes 30 (50.0) |
| Channel subscribers, n       | Median (IQR) 19950 (3370–50900) |
|                              | Range 0–894000 |
| Video production country, n (%) | India 45 (75.0) |
|                              | Rest of the World 15 (25.0%) |
| Video author, n (%)          | Healthcare providers 45 (75.0) |
|                              | Patients 12 (20.0) |
|                              | Other 3 (5.0) |
| Video target, n (%)          | Healthcare providers 16 (26.7) |
|                              | Patients 44 (73.3) |
| Video topic, n (%)           | Medical lesson 28 (46.7) |
|                              | Information for patients 18 (30.0) |
|                              | Patient’s experience 12 (20.0) |
|                              | Commercials 2 (3.3) |
| Video uploaded year, n (%)   | 2013–2017 16 (26.7) |
|                              | 2018–2020 44 (73.3) |

Abbreviations: IQR Interquartile Range.

**Video quality assessment**

The Cohen kappa statistic was used to measure the reliability of the investigator’s assessments between the two evaluation times and ranged from 0.61 (PEMAT A/V Understandability) to 1.0 (for Misinformation item 1), indicating a substantial and almost perfect agreement.

The median PEMAT Understandability score was 42.9% (IQR:34.5–58.9), and the median PEMAT Actionability score was 50.0% (IQR:25.0–56.2, Table 2).

When PEMAT A/V assessment tool was examined over time, the median Understandability score ranged from 28.6% in 2013 to 42.9% in 2020 (p = 0.08, Fig. 2A) and the median Actionability score ranged from 25.0% in 2013 to 50.0% in 2020 (p = 0.5, Fig. 2B).

**Table 2.** PEMAT audio/visual (A/V) Understandability and Actionability of 60 YouTube™ videos on “phimosis” selected on January 16th, 2021.

| PEMAT Understandability | Mean (SD) |
|-------------------------|-----------|
| Median (IQR)            | 45.8 (2.5) |
| PEMAT Actionability     | Mean (SD) |
| Median (IQR)            | 45.4 (3.3) |
| PEMAT Actionability     | Median (IQR) |
|----------               | 50.0 (25.0–56.2) |

The Patient Education Materials Assessment Tool (PEMAT) for audio/visual materials (A/V) is a systematic method consisting of 13 items measuring Understandability and 4 items measuring Actionability. Abbreviations: SD Standard Deviation; IQR Inter Quartile Range; PEMAT Patient Educational Assessment Tool.

The median misinformation score ranged from 2.0 (item 3: “Provides correct explanation of the surgical option” and item 5: “Presents and explains treatment results clearly”) to 4.0 (item 1: “Accurately explains the topic”; item 2: “Factually describes when the condition is pathological”). According to target videos (patients vs healthcare providers), we recorded a statistically significant difference in item 3 “Provides correct explanation of the surgical option” (median 2.0 [IQR:0.0–3.0] vs 4.0 [IQR:2.0–4.2], p ≤ 0.001) and in item 4 “Mentions the EAU approved recommendations” (median 2.0 [IQR:1.0–4.0] vs 3.0 [IQR:3.0–4.0], p < 0.01, Table 3). When the median score of the misinformation scale was examined over time, although it varied some over the past 8 years, it remained 2.6 in 2013 and in 2020 (p = 0.1, Fig. 2C).

**Variable correlations**

We tested for possible correlations between video-graphic characteristics (length, thumbs-up, thumbs-down, number of comments, persistence on YouTube™, channel subscribers) and quality assessment tools (PEMAT Understandability and Actionability, and misinformation scale). Specifically, we recorded a statistically significant positive correlation between thumbs-up and Understandability (r = 0.44, p < 0.001), number of comments and Understandability (r = 0.41, p = 0.001), channel subscribers and Actionability (r = 0.32, p = 0.02), and length and misinformation scale (r = 0.39, p < 0.01). Conversely, no statistically significant correlations were found between other comparisons (all p > 0.05).

Finally, we tested for a possible correlation between PEMAT Understandability, PEMAT Actionability, and the misinformation scale. Specifically, we recorded a statistically significant positive correlation between PEMAT Understandability and misinformation scale (r = 0.44, p < 0.001), and PEMAT Actionability and the misinformation scale (r = 0.35, p < 0.01).

**DISCUSSION**

The current study aims to evaluate the quality of information provided in YouTube™ videos on phimosis. To the best of our knowledge, no previous research has examined this topic, and our study identified several noteworthy observations.

First, out of 60 videos, three-quarters were uploaded by healthcare providers and approximately half were patient-targeted. Moreover, we recorded that in 50% of videos, the users take advantage of comments to ask for medical advice. These findings highlight that YouTube™ may be used by users to look and ask for information on medical issues, in particular on phimosis, instead of seeking a urological consultation. Therefore, it is important to provide high-quality information for patients to avoid mistreatment and the delay of specialist consultation. For example, this phenomenon may induce an increase in self-circumcision practice, an emerging phenomenon [24, 25]. In this context, misleading information may lead to a higher risk of complications, such as foreskin necrosis [26].
Secondly, we discovered that out of 60 videos, 75% were produced in India, a country with a high impact of penile cancer (incidence rate of up to 3.32 per 100,000 men in some regions) [27]. The mortality related to penile cancer ranges from 20% to 91% according to the cancer stage [28]. It is known that phimosis represents a risk factor for penile squamous cell carcinoma [29]. This observation may explain the high amount of YouTube™ videos on phimosis in this country. However, the rest of the world should not underestimate this medical issue and should aim to increase both the volume and quality of the internet content on phimosis.

Third, according to the PEMAT A/V tool, median Understandability score was 42.9% and the median Actionability was 50.0%. According to Shoemaker et al. [21], videos with PEMAT Understandability and Actionability scores of <70% are considered poorly understandable or poorly actionable. Therefore, the scores recorded in the current analysis were low. The Understandability score reflects how viewers, regardless of cultural backgrounds and medical knowledge, could process and explain key messages provided. The Actionability score reflects how viewers could use the information presented. Thus, the YouTube™ videos neither served an educational role in disseminating quality phimosis information nor could patients easily use the information provided.

To the best of our knowledge, we are the first to analyze the quality of YouTube™ videos information on phimosis. Similar findings were observed in other studies dealing with the fairness of YouTube™ videos on different medical diseases. For instance, Salama et al. analyzed the quality of videos on hypospadias, reporting Understandability and Actionability scores of 54.5% and 21.8%, respectively [30]. In Rubel et al. study, Understandability and Actionability scores of 46.3% and 57.7%, respectively, were recorded in the analysis of YouTube™ videos on sinusitis [31]. In conclusion, similarly to Salama et al. and Rubel et al., the PEMAT A/V scores recorded in our analysis were also low. Nonetheless, we recorded an improvement of PEMAT A/V Understandability (Δ = 14.3%) and Actionability (Δ = 25.0%) in the most recently uploaded videos, with respect to the past. These observations may be interpreted as a sign of increasing awareness of the YouTube™ videos’ utility as a source for mass information. However, further improvement is needed to ensure higher quality of information on phimosis can be accessed by users.

According to the misinformation scale, we recorded the lowest median score for item 3 “Provides correct explanation of the surgical option” and item 5 “Presents and explains treatment results clearly”. Conversely, the highest median score was recorded for item 1 “Accurately explains the topic” and item 2 “Factually describes when the condition is pathological”. As a consequence, users accessing YouTube™ videos were not well-informed on surgical options and treatment results, but they were aware of their condition. Moreover, according to the video target audience, videos targeted towards healthcare providers had higher scores for items 3 “Provides correct explanation of the surgical option” and 4 “Mentions the EAU approved recommendations”, with respect to videos targeted to patients. In conclusion, these considerations...
emphasized that authors of YouTube™ videos provided lower quality information when the patients represented the target audience rather than healthcare providers. Moreover, different from PEMAT A/V, no improvement in the misinformation score was observed.

Finally, a Pearson’s correlation between video-graphic content and quality assessment tools was tested. This correlation revealed that videos with higher number of thumbs-up and comments had better Understandability (r = 0.44, p < 0.001; r = 0.41, p = 0.001, respectively). Consequently, the more understandable the video was, the more patients were encouraged to click thumbs-up and leave a comment. Additionally, videos posted by channels with higher number of subscribers had better Actionability (r = 0.32, p = 0.02), indicating that the better the videos explained how to make decisions, the more users were encouraged to follow the channel. Regarding the length of videos, we recorded that videos with longer lengths had higher misinformation scores (r = 0.39, p < 0.01). Thus, when the video-makers spent more time explaining their content, it was more likely that their content was of higher quality, and they were less likely to provide incomplete information or misinformation. Lastly, videos with higher PEMAT Understandability and PEMAT Actionability had better misinformation scores (r = 0.44, p < 0.001; r = 0.35, p < 0.01, respectively). Therefore, the more understandable and actionable the videos were, the higher the quality of information that was provided to YouTube™ users.

Taken together, according to the analysis that we conducted using validated instruments, we can deduce that the quality of the information uploaded on YouTube™ videos is low but it is increasing in the last years. Thus, at present YouTube™ cannot be recommended as a source of medical information on this condition. Similar considerations arise from previous studies on other urological issues, such as erectile dysfunction [14], premature ejaculation [13], testicular cancer [32], hypospadias [30], urinary stones [33], penile prosthesis [34], testicular pain [35], bladder pain syndrome [36], and Peyronie’s disease [37]. However, since we are living in a digital era and because of the increasing use of telemedicine, especially during the Covid 19 pandemic, patients are looking for medical advice on the Internet [9, 38]. For these reasons, it would be beneficial for trusted medical institutions and associations to make an effort to offer high-quality, medical content on YouTube™, in order to provide good-quality content to the general public.

Our study is not devoid of limitations. First, YouTube™ search results rely on Google’s proprietary search algorithms based on users’ previous search activities and location. To minimize these effects, before searching, a log out from any personal accounts and a change of location via VPN proxy, were performed. Second, only English-language videos were included; videos in other languages may provide different results. Third, some reliable videos might be missed, due to our search terms. However, we assumed that uploaders generally intend to use the word “phimosis” either in the title or keywords. Fourth, the quality assessment of videos was a subjective evaluation. However, to reduce this confounder, three investigators were involved to independently analyze video contents. Finally, our analysis was limited, according to the selection criteria, to 60 videos. However, this sample reflects what the majority of the general population views on Internet resources (only the first 3 pages of content, 60 videos) [19]. Regardless of these limitations, the present study can be considered as a snapshot of the current phimosis’ information in YouTube™ videos.

In conclusion, the overall quality of the information on phimosis uploaded on YouTube™ videos is low. A trend toward an improvement of PEMAT A/V Understandability and Actionability over time is evident, but not statistically significant. Therefore, at the moment, YouTube™ cannot be recommended as a reliable source of information on phimosis. Video producers should upload more accurate, higher-quality videos to help physicians and patients during the decision-making process.

DATA AVAILABILITY
Data and materials are available whenever requested.

CODE AVAILABILITY
Code is available whenever requested.

REFERENCES
1. Salonia A, Bettocchi C, Boeri L, Capogrosso P, Carvalho J, Cilesiz NC, et al. European association of urology guidelines on sexual and reproductive health—2021 update: male sexual dysfunction. Eur Urol. 2021;80:333–57.
2. Hayashi Y, Koijima Y, Mizuno K, Kohri K. Precipuce: phimosis, paraphimosis, and circumcision. Sci World J. 2011;11:289–301.

Table 3. Misinformation scale evaluating 60 YouTube™ videos on “phimosis”, selected on January 16th, 2021, according to target video (Patients vs Healthcare providers).

| Misinformation scale                                                                 | Overall videos (n = 60) | Patients (n = 44, 73.3%) | Healthcare providers (n = 16, 26.7%) | p-value |
|--------------------------------------------------------------------------------------|------------------------|-------------------------|-------------------------------------|---------|
| 1. Accurately explains the topic                                                     | Mean (SD) 3.1 (0.2)    | 3.1 (0.3)               | 3.1 (0.3)                           | 0.9     |
| 2. Factualy describes when the condition is pathological                            | Mean (SD) 3.0 (0.2)    | 3.0 (0.2)               | 3.2 (0.3)                           | 0.7     |
| 3. Provides correct explanation of the surgical option                               | Mean (SD) 2.4 (0.2)    | 2.0 (0.2)               | 3.4 (0.4)                           | 0.001   |
| 4. Mentions the EAU approved recommendations                                         | Mean (SD) 2.7 (0.2)    | 2.4 (0.2)               | 3.4 (0.2)                           | <0.01   |
| 5. Presents and explains treatment results clearly                                    | Mean (SD) 2.1 (0.2)    | 2.0 (0.2)               | 2.6 (0.4)                           | 0.1     |
| Misinformation score                                                                | Median (IQR) 3.0 (2.0–4.0) | 2.0 (1.0–4.0)         | 3.0 (3.0–4.0)                       | <0.01   |

A 5-item Misinformation scale was created according to EAU Patient Information. Each item was rated from 1 to 5 (1-Strongly Disagree; 2-Disagree; 3-Neither agree nor disagree; 4-Agree; 5-Strongly Agree).

Abbreviations: SD Standard Deviation; IQR Interquartile Range; EAU European Association of Urology.
28. Survival Rates for Penile Cancer [Internet]. [cited 2021 Dec 20]. Available from: https://www.cancer.org/cancer/penile-cancer/detection-diagnosis-staging/survival-rates.html.

ACKNOWLEDGEMENTS

The authors thank Rachel C. Applefield, from the IRCSS San Raffaele Scientific Institute, for the English language revision of this paper. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS

Conception and design of the study: SC, CCR, MC, MG, RLR, GC, GC. Data acquisition: JC, SC, RP, MRC, VT, TN, MV, SF. Analysis and interpretation of data: SC, RLR, GC, SG. Writing the manuscript: AP, CI, VM. Style revision: FC, PV. All authors revised the manuscript and approved the version submitted.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL

The paper is exempt from ethical committee approval.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Claudia Collà Ruvolo.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.