

Systematic analysis of the scientific quality, reliability, and educational value of Vaginal Natural Orifice Transluminal Endoscopic Surgery videos on YouTube

Okay Alptekin^{1*} , Betül Kalkan² 

SUMMARY

OBJECTIVE: The aim of this study was to evaluate the scientific quality, reliability, and educational value of YouTube videos on Vaginal Natural Orifice Transluminal Endoscopic Surgery using structured assessment tools.

METHODS: This cross-sectional study analyzed 20 YouTube videos on Vaginal Natural Orifice Transluminal Endoscopic Surgery identified using the keywords "Vaginal Natural Orifice Transluminal Endoscopic Surgery," "Vaginal Natural Orifice Transluminal Endoscopic Surgery hysterectomy," "vaginal natural orifice endoscopic surgery," and "transvaginal NOTES." Two obstetrics and gynecology specialists independently evaluated the videos using the Global Quality Scale, modified DISCERN, and Journal of the American Medical Association benchmarks. Analyses included non-parametric group comparisons, Spearman correlation, multiple linear regression, and intraclass correlation coefficient analysis.

RESULTS: The median number of views was 2,565 (interquartile range: 13,322). The mean Global Quality Scale, modified DISCERN, and Journal of the American Medical Association scores were 4.17 ± 0.94 , 3.95 ± 0.84 , and 2.65 ± 0.93 , respectively. Inter-rater agreement was moderate for Global Quality Scale and modified DISCERN and good for Journal of the American Medical Association. A statistically significant difference by channel type was observed only for modified DISCERN scores. Although Global Quality Scale scores were moderately correlated with likes ($p=0.565$, $p=0.009$), multiple linear regression showed no independent association between popularity or technical quality variables and quality scores ($p>0.05$).

CONCLUSION: Vaginal Natural Orifice Transluminal Endoscopic Surgery videos on YouTube may provide valuable educational content; however, popularity metrics do not reflect scientific reliability. Video selection for surgical education should be based on structured quality assessment tools rather than view count alone.

KEYWORDS: Natural orifice endoscopic surgery. Social media. Video recording. Education, medical. Gynecology. Reproducibility of results.

INTRODUCTION

Vaginal Natural Orifice Transluminal Endoscopic Surgery (vNOTES) provides vaginal access to the peritoneal cavity by combining endoscopic and laparoscopic methods, thereby avoiding an abdominal skin incision and potentially reducing perioperative morbidity¹. It has been adopted as an alternative to conventional laparoscopy or open surgery in benign gynecologic procedures such as hysterectomy and adnexal surgery². However, limited vaginal access, previous pelvic surgery, and advanced pelvic adhesions may restrict its applicability and require careful patient selection³.

Surgical education has evolved from the traditional master-apprentice model toward digital learning environments, and YouTube has become one of the most frequently used platforms for health-related and surgical educational content because of

its open-access nature⁴. However, the absence of peer review raises concerns regarding accuracy: health-related and surgical videos are frequently incomplete, misleading, or heterogeneous in educational structure, presentation of technical steps, and source transparency^{5,6}. Structured scoring systems have therefore been recommended to assess such material reliably⁷, and popularity metrics have not been found to reflect scientific quality across surgical fields^{8,9}.

The literature has nevertheless focused mainly on common procedures, and studies evaluating YouTube content for relatively novel, technically demanding procedures with a pronounced learning curve, such as vNOTES, remain limited. vNOTES has a distinct learning curve due to its different anatomical orientation, limited instrument triangulation, and technical constraints in vaginal access¹⁰.

¹Erzincan Mengücek Gazi Training and Research Hospital, Department of Obstetrics and Gynecology – Erzincan, Turkey.

²Erzincan Binali Yıldırım University, Faculty of Medicine, Department of Obstetrics and Gynecology – Erzincan, Turkey.

*Corresponding author: okayalptekin@gmail.com

Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.

Received on July 03, 2026. Accepted on August 08, 2026.

Scientific Editor: José Maria Soares Júnior 

Studies evaluating the educational value of vNOTES videos remain limited, and the available evidence suggests that such content cannot be considered a reliable standalone resource for surgical education¹¹. The present study differs from earlier YouTube analyses in three respects. First, it addresses a procedure still in its diffusion phase, for which structured proctorship is scarce, and trainees therefore rely more heavily on freely available video material. Second, educational quality, scientific reliability, and source transparency were assessed simultaneously with three complementary instruments by two independent and mutually blinded evaluators, with inter-rater agreement quantified rather than assumed. Third, the assumption that popularity indicates quality was tested in a multivariable model rather than by correlation alone. The aim of this study was to evaluate the scientific quality, reliability, and educational value of vNOTES videos on YouTube using structured assessment tools.

METHODS

Video screening strategy

This cross-sectional study was conducted in January 2026 and analyzed YouTube videos using the keywords “vNOTES surgery,” “vNOTES hysterectomy,” “vaginal natural orifice endoscopic surgery,” and “transvaginal NOTES.” Searches were performed in incognito mode to minimize bias. The first 50 results sorted by relevance for each keyword were screened. After duplicate removal, the remaining videos were assessed according to inclusion and exclusion criteria. Video selection was performed independently by two researchers, and disagreements were resolved by consensus.

Inclusion and exclusion criteria

The study included videos directly related to vNOTES surgery that contained narration or subtitles in English or Turkish, allowed visualization of surgical steps, had a duration of at least 1 min, more than 100 views, and adequate image quality to clearly identify the surgical field (minimum 720p). Because vNOTES is a relatively recent surgical technique and the amount of available content is limited, a view-count threshold of more than 100 views was applied so that technically relevant videos would not be excluded.

Content intended for advertising or promotion, videos containing only patient experiences or interviews, theoretical presentations without surgical procedures, duplicate videos, and low-quality videos that were silent, lacked subtitles, or did not clearly show the surgical field were excluded. A total of 20 videos were included in the final analysis.

Data collection and assessment scales

For each video, the following data were recorded: duration, upload date, views, likes, comments, language, channel type (academic journal, educational institution, or individual/medical channel), and whether source or reference information was provided in the description section. Data were collected using a standardized form. Video screening and scoring were performed by the two evaluators independently. Each evaluator was blinded to the other's scores, and no scores were shared or discussed until all evaluations were completed and the intra-class correlation coefficients had been calculated.

Both evaluators were blinded to video popularity metrics and scored the videos on content alone. The Global Quality Scale (GQS) assessed overall educational quality on a 1–5 scale, with higher scores indicating better quality.

The modified DISCERN (mDISCERN) scale assessed scientific reliability through five criteria: source citation, impartial presentation, consistency with current scientific evidence, discussion of risks and alternatives, and clarity; each was scored 1 for “yes,” giving a total of 0 to 5.

The Journal of the American Medical Association (JAMA) criteria assessed transparency and academic accountability through authorship, source attribution, disclosure of conflicts of interest or sponsorship, and content currency, with one point per criterion and a total of 0 to 4.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics version 27 (IBM Corp., Armonk, NY, USA). Continuous variables were assessed using the Shapiro-Wilk test and were not normally distributed ($p < 0.05$ for all variables). Therefore, data were presented as mean $\pm$ standard deviation and median with interquartile range (IQR), and non-parametric tests were used for group comparisons.

The Kruskal-Wallis test was used to compare quality scores by channel type, and significant differences were examined with the Mann-Whitney U test. Relationships between continuous variables were assessed using Spearman's rank correlation coefficient (ρ).

Inter-rater agreement was assessed using the single-measure intraclass correlation coefficient (ICC [2,1]) based on a two-way random-effects model with absolute agreement and reported with 95%CI. ICC values were interpreted as <0.50 , poor; 0.50 – 0.75 , moderate; 0.75 – 0.90 , good; and >0.90 , excellent.

Multiple linear regression was performed with view count, like count, and video quality as independent variables and GQS, mDISCERN, and JAMA scores as separate dependent variables; multicollinearity, assessed by variance inflation factor, was within acceptable limits.

Two-sided p-values were used in all statistical analyses, and $p < 0.05$ was considered statistically significant.

Ethics statement

Ethics committee approval was obtained from the Erzincan Binali Yildirim University Non-Interventional Clinical Research Ethics Committee (decision no: 2025-18/06; date: 16.10.2025). The study followed the Declaration of Helsinki and national ethical standards. As it analyzed only publicly available YouTube videos and involved no direct human participation or identifiable personal data, informed consent was not required.

RESULTS

Descriptive statistics and video characteristics

The main characteristics of the 20 included videos are presented in Table 1. The median video duration was 282.5 s (IQR: 290.8; range: 81–1920), the median view count was 2,565 (IQR: 13,322; range: 168–56,650), the median number of likes was 25.5 (IQR: 67.8), and the median number of comments was 0 (IQR: 1.0).

Mean quality scores were 4.17 ± 0.94 for the Global Quality Scale (GQS), 3.95 ± 0.84 for mDISCERN, and 2.65 ± 0.93 for the JAMA criteria (Figure 1).

Seventy-five percent of the videos ($n=15$) were in English, and 25% ($n=5$) were in Turkish. According to channel type,

educational channels constituted the largest group at 55% ($n=11$), followed by individual medical channels at 30% ($n=6$) and academic journal-based channels at 15% ($n=3$).

Inter-rater reliability (intraclass correlation coefficient)

In the inter-rater reliability analysis between the two independent raters, good reliability was observed for the JAMA scale (ICC=0.706; 95%CI 0.40–0.87; $p < 0.001$). Moderate reliability was observed for the GQS (ICC=0.609; 95%CI 0.25–0.82; $p=0.001$) and mDISCERN (ICC=0.518; 95%CI 0.12–0.78; $p=0.007$) scales (Table 2).

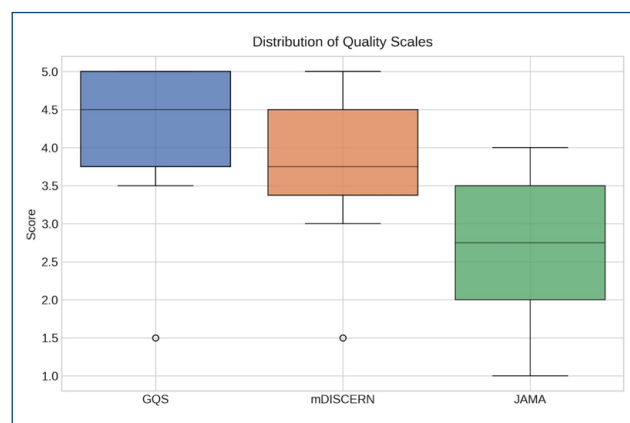


Figure 1. Distribution of Global Quality Scale, modified DISCERN, and Journal of the American Medical Association scores (box plots).

Table 1. General characteristics and descriptive statistics of the videos.

Variable	Mean±SD	Median (IQR)	Min–max
Duration (seconds)	433.2±470.5	282.5 (290.8)	81–1,920
Views	9,095±13,551	2,565 (13,322)	168–56,650
Likes	41.1±45.5	25.5 (67.8)	3–185
Comments	1.1±2.0	0.0 (1.0)	0–6
Mean GQS score	4.17±0.94	4.50 (1.5)	1.5–5.0
Mean mDISCERN score	3.95±0.84	4.00 (1.0)	1.5–5.0
Mean JAMA score	2.65±0.93	2.75 (1.5)	1.0–4.0

SD: standard deviation; IQR: interquartile range; GQS: Global Quality Scale; mDISCERN: modified DISCERN; JAMA: Journal of the American Medical Association.

Table 2. Inter-rater reliability analysis (intraclass correlation coefficient).

Scale	ICC (single absolute)	95%CI	p-value	Level of agreement
GQS	0.609	0.25–0.82	0.001	Moderate
mDISCERN	0.518	0.12–0.78	0.007	Moderate
JAMA	0.706	0.40–0.87	<0.001	Good

ICC: intraclass correlation coefficient; CI: confidence interval; GQS: Global Quality Scale; mDISCERN: modified DISCERN; JAMA: Journal of the American Medical Association.

Comparisons by channel type and language

A statistically significant difference by channel type was observed only for mDISCERN scores (Kruskal-Wallis, $p=0.041$). Journal-based channels had higher mDISCERN scores than educational and medical channels. However, this group comprised only three videos, and this finding should therefore be interpreted with caution. No significant differences were found for GQS ($p=0.083$) or JAMA scores ($p=0.483$). No statistically significant differences were found between Turkish and English videos for GQS, mDISCERN, or JAMA scores ($p>0.05$ for all comparisons).

Correlation and regression analyses

In the Spearman correlation analysis, a moderate positive correlation was found between GQS scores and views ($p=0.490$, $p=0.028$) and likes ($p=0.565$, $p=0.009$). In contrast, no significant relationship was found between mDISCERN and JAMA scores and views, likes, comments, or video quality ($p>0.05$). In multiple linear regression analysis, none of the models, including view count, like count, and video quality, were significantly associated with GQS, mDISCERN, or JAMA scores. The GQS ($R^2=0.312$, $p=0.104$), mDISCERN ($R^2=0.238$, $p=0.214$), and JAMA models were all non-significant, and only overall model fit is therefore reported.

DISCUSSION

The findings showed that vNOTES-related YouTube videos generally had good educational quality (GQS: 4.17 ± 0.94) and satisfactory scientific reliability (mDISCERN: 3.95 ± 0.84), but only moderate transparency (JAMA: 2.65 ± 0.93). These findings differ from systematic reviews describing surgical YouTube videos as low quality. Gorgy et al. reported that 96.6% of surgical videos were educationally unsatisfactory⁹, and health-related YouTube content in general has been judged to remain below acceptable standards⁴. In the only previous vNOTES-specific analysis, Temur reported that 43.1% of videos were low quality and that the median GQS score was 2¹¹.

This may be explained by the inclusion criteria applied in our study, which required a minimum resolution of 720p, a minimum duration of 1 min, and clear visualization of the surgical field. These criteria may have improved the overall quality of the sample by excluding low-resolution or amateur content that could contribute to misinformation. In addition, 15% of the videos were sourced from academic journals and 55% from structured educational channels, which may have contributed to the higher reliability observed in journal-based channels ($p=0.041$). However, only three videos originated from

journal-based channels, and this observation should therefore be regarded as preliminary. Comparable observations have been reported for other gynecologic procedures, where videos uploaded predominantly by healthcare professionals nevertheless achieved only moderate overall quality and low mDISCERN scores because of insufficient academic referencing^{12,13}. These findings suggest that high visual quality, reflected by GQS, does not always correspond to high scientific reliability, reflected by mDISCERN.

Videos produced by academic institutions, professional organizations, or surgeons have consistently been reported to provide higher educational value and transparency than those from individual users^{6,9,11}.

One of the main findings of our study was that view and like counts had no independent effect on quality scores in multiple linear regression analysis. Assuming that the most viewed video is also the most reliable may therefore be misleading, as platform algorithms prioritize engagement over accuracy and allow scientifically flawed techniques to reach large audiences⁴. Therefore, structured assessment tools such as GQS, mDISCERN, and JAMA should be prioritized over popularity metrics in surgical educational materials.

vNOTES has a distinct learning curve because of its reverse anatomical perspective and limited instrument movement, with approximately 16 cases required for operative stability and 18 for proficiency in vNOTES hysterectomy¹⁰. However, Kapurubandara et al. emphasized that attempting innovative techniques such as vNOTES based solely on YouTube videos may involve serious surgical risks without structured proctorship training³. The educational value of video materials depends on clear demonstration of surgical safety, including the Critical View of Safety; however, compliance with these standards has generally been reported to be below 10%⁹.

The study was limited by the small sample of 20 videos, which restricts generalizability, and by the dynamic nature of YouTube content, meaning that the data reflect a single time point¹¹. The sample was small because the number of vNOTES videos meeting the predefined technical and content criteria was limited; findings based on 20 videos, and in particular subgroup comparisons involving as few as three videos, should therefore be regarded as descriptive and hypothesis-generating rather than definitive. In addition, the inclusion criteria applied here (minimum 720p resolution, minimum 1-min duration, and clear visualization of surgical steps) are likely to have preferentially selected structured, institutional, or professionally produced content. The analyzed sample therefore represents a preselected, comparatively high-quality subset rather than the full spectrum of vNOTES content on YouTube. This may explain the higher

scores observed compared with broader YouTube analyses and limits the generalizability of our findings. Peer-reviewed official video channels created by surgical societies and academic institutions may improve digital medical education and help counter anecdotal bias on YouTube⁴. For surgical education, these findings support a practical approach: video material should be selected by structured criteria rather than view count, preferably from accredited society or journal channels; it should complement rather than replace proctored hands-on training; and trainees should verify authorship, referencing, and disclosure before adopting a technique observed online. Curated, peer-reviewed vNOTES video libraries in which safety-critical steps are explicitly demonstrated would support this.

CONCLUSION

Within this preselected sample, YouTube videos on vNOTES provided visually adequate educational content, although quality varied considerably. Videos from academic journal-based sources tended to show the highest reliability scores; however, this group comprised only three videos and the observation requires confirmation in larger samples. Because view and like

counts are insufficient indicators of quality, structured assessment tools such as GQS and DISCERN should be used in selecting videos for surgical education. As our inclusion criteria yielded a technically preselected subset, these findings cannot be generalized to all vNOTES content on YouTube. Accredited channels endorsed by surgical societies may further improve the reliability of digital medical education.

AUTHORS' CONTRIBUTIONS

OA: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **BK:** Funding acquisition, Investigation, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

REFERENCES

- Wylie S, Mutema E, Yoong W. Vaginal natural orifice transluminal endoscopic surgery (vNOTES): a new chapter in vaginal surgical innovation. *J Obstet Gynaecol*. 2025;45(1):2458842. <https://doi.org/10.1080/01443615.2025.2458842>
- Marchand GJ, Ulibarri H, Arroyo A, Blanco M, Herrera DG, Hamilton B, et al. Systematic review and meta-analysis of vaginal natural orifice transluminal endoscopic surgery hysterectomy versus vaginal hysterectomy for benign indications. *AJOG Glob Rep*. 2024;4(2):100355. <https://doi.org/10.1016/j.xagr.2024.100355>
- Kapurubandara S, Lowenstein L, Salvay H, Herijgers A, King J, Baekelandt J. Consensus on safe implementation of vaginal natural orifice transluminal endoscopic surgery (vNOTES). *Eur J Obstet Gynecol Reprod Biol*. 2021;263:216–22. <https://doi.org/10.1016/j.ejogrb.2021.06.019>
- Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med Educ*. 2022;22(1):382. <https://doi.org/10.1186/s12909-022-03446-z>
- Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. *Health Informatics J*. 2015;21(3):173–94. <https://doi.org/10.1177/1460458213512220>
- Noll D, Green L, Asokan G, Barbaro A, Hewitt J, Ovenden C, et al. Is YouTube™ a good source of information for patients to understand laparoscopic fundoplication? *Ann R Coll Surg Engl*. 2023;105(4):365–71.
- de'Angelis N, Gavrilidis P, Martínez-Pérez A, Genova P, Notarnicola M, Reitano E, et al. Educational value of surgical videos on YouTube: quality assessment of laparoscopic appendectomy videos by senior surgeons vs. novice trainees. *World J Emerg Surg*. 2019;14:22. <https://doi.org/10.1186/s13017-019-0241-6>
- Erdem MN, Karaca S. Evaluating the accuracy and quality of the information in kyphosis videos shared on YouTube. *Spine (Phila Pa 1976)*. 2018;43(22):E1334–9. <https://doi.org/10.1097/brs.0000000000002691>
- Gorgy A, Hawary HE, Galli R, MacDonald M, Barone N, Thibaudeau S. Evaluating the educational quality of surgical YouTube® videos: a systematic review. *Health Sci Rev*. 2022;5:100067. <https://doi.org/10.1016/j.hsr.2022.100067>
- Housmans S, Baekelandt J, Deprest J. Proctored single surgeon learning curve for vaginal natural orifice transluminal endoscopic surgery (NOTES) hysterectomy. *Eur J Obstet Gynecol Reprod Biol*. 2025;307:223–9. <https://doi.org/10.1016/j.ejogrb.2025.02.020>
- Temur I. Evaluation of the educational value, reliability, and quality of vaginally-assisted natural orifice transluminal endoscopic surgery (v-NOTES) videos on YouTube. *Cureus*. 2025;17(6):e85244. <https://doi.org/10.7759/cureus.85244>
- Erdogan G. Female genital cosmetic surgery (FGCS): evaluation of YouTube videos. *J Gynecol Obstet Hum Reprod*. 2021;50(4):102102. <https://doi.org/10.1016/j.jogoh.2021.102102>
- Özcan C, Sancı A. Evaluation of robotic-assisted sacrocolpopexy videos on YouTube. *Int Urogynecol J*. 2023;34(8):1843–7. <https://doi.org/10.1007/s00192-023-05480-x>

