

The quality and reliability of short videos about Premature Ovarian Failure on BiliBili and TikTok: Cross-sectional study

Ren Xu, Yanan Ren, Xinjun Li, Yanmei Sun, Luyang Su, Jianzhi Su

Submitted to: JMIR Formative Research
on: March 24, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	21

Preprint
JMIR Publications

The quality and reliability of short videos about Premature Ovarian Failure on Bilibili and TikTok: Cross-sectional study

Ren Xu¹; Yanan Ren¹; Xinjun Li¹; Yanmei Sun¹; Luyang Su¹; Jianzhi Su²

¹ Hebei General Hospital shijiazhuang CN

² The Fourth Hospital of Hebei Medical University shijiazhuang CN

Corresponding Author:

Jianzhi Su

The Fourth Hospital of Hebei Medical University
No. 12, Health Road, Shijiazhuang, Hebei Province, China
shijiazhuang
CN

Abstract

Background: Premature Ovarian Failure (POF) is a condition where ovarian function ceases before 40, causing infertility and other health issues. With growing public health awareness, platforms like TikTok and Bilibili are key sources of health information, but concerns about content quality and reliability remain.

Objective: his study assesses the quality of POF-related videos on these platforms.

Methods: A cross-sectional study was conducted, analyzing 187 POF-related videos from TikTok and Bilibili. The modified DISCERN and Global Quality Score (GQS) assessment tools were used to evaluate video reliability and quality. Data on video characteristics, including engagement metrics and content, were also collected. Statistical analyses were performed to assess differences between platforms, video sources, and video quality.

Results: Result?TikTok is more popular than Bilibili. Videos on both platforms related to POF had suboptimal quality, with TikTok's GQS median at 3 (2-3) and Bilibili's at 2 (1-3), showing a significant difference ($P < 0.05$). However, both had median modified DISCERN scores of 2 (1-3), with no significant difference ($P > 0.05$). On Bilibili, video duration was positively correlated with quality scores ($P < 0.001$), but no such correlation was found on TikTok. Symptoms of POF were the most common topic, with 29% of videos providing detailed explanations. Diagnosis and treatment were discussed in 16% and 7.4% of videos, respectively. Expert-uploaded videos were of significantly higher quality than those from non-experts and personal users, with experts showing higher GQS and modified DISCERN scores.

Conclusions: ikTok had higher engagement and better quality than Bilibili, but both platforms had inadequate video quality and reliability on POF. Expert videos were more reliable. These findings highlight the need for better regulation and monitoring of health content on short-video platforms.

(JMIR Preprints 24/03/2025:74416)

DOI: <https://doi.org/10.2196/preprints.74416>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in [http](#)

No. Please do not make my accepted manuscript PDF available to anyone.



Original Manuscript

The quality and reliability of short videos about Premature Ovarian Failure on Bilibili and TikTok: Cross-sectional study

Ren Xu¹, Yanan Ren¹, Xinjun Li¹, Yamei Sun², Luyang Su³, Jianzhi Su^{4*}

¹Department of Gynecology, Hebei General Hospital, Shijiazhuang, Hebei-China.

²Department of Reproductive Genetics, Hebei General Hospital, Shijiazhuang, Hebei-China.

³Physical Examination Center, Hebei General Hospital, Shijiazhuang, Hebei-China.

⁴Department of Urology Surgery, The Fourth Hospital of Hebei Medical University, Shijiazhuang, Hebei-China

* Correspondence:

*Jianzhi Su

E-mail:18533167591@hebmu.edu.cn

Abstract

Background: Premature Ovarian Failure (POF) is a condition where ovarian function ceases before 40, causing infertility and other health issues. With growing public health awareness, platforms like TikTok and Bilibili are key sources of health information, but concerns about content quality and reliability remain. This study assesses the

quality of POF-related videos on these platforms.

Methods: A cross-sectional study was conducted, analyzing 187 POF-related videos from TikTok and Bilibili. The modified DISCERN and Global Quality Score (GQS) assessment tools were used to evaluate video reliability and quality. Data on video characteristics, including engagement metrics and content, were also collected. Statistical analyses were performed to assess differences between platforms, video sources, and video quality.

Result □ TikTok is more popular than Bilibili. Videos on both platforms related to POF had suboptimal quality, with TikTok's GQS median at 3 (2-3) and Bilibili's at 2 (1-3), showing a significant difference ($P < 0.05$). However, both had median modified DISCERN scores of 2 (1-3), with no significant difference ($P > 0.05$). On Bilibili, video duration was positively correlated with quality scores ($P < 0.001$), but no such correlation was found on TikTok. Symptoms of POF were the most common topic, with 29% of videos providing detailed explanations. Diagnosis and treatment were discussed in 16% and 7.4% of videos, respectively. Expert-uploaded videos were of significantly higher quality than those from non-experts and personal users, with experts showing higher GQS and modified DISCERN scores.

Conclusions: TikTok had higher engagement and better quality than Bilibili, but both platforms had inadequate video quality and reliability on POF. Expert videos were more reliable. These findings highlight the need for better regulation and monitoring of health content on short-video platforms.

Key words: Premature Ovarian Failure; Short Videos; Information Quality; Social Media; TikTok; Bilibili; Global Quality Score; Modified DISCERN; Reliability

Introduction

Premature Ovarian Failure (POF) refers to the loss of ovarian function before the age of 40, resulting in the cessation of menstruation and infertility, with an increasing prevalence in recent years¹. Beyond fertility issues, POF is associated with menopausal symptoms, such as hot flashes, osteoporosis, and an elevated risk of cardiovascular diseases, all of which significantly impact the overall health and quality of life of affected individuals^{2, 3}. Additionally, the onset of POF is often accompanied by psychological health issues, including anxiety and depression, making it a multidimensional public health challenge^{4, 5}. Therefore, early identification, intervention, and raising public awareness about POF are crucial for improving patient health and societal well-being.

In the digital age, the way the public accesses health information has undergone a significant transformation^{6, 7}. While traditional textual and academic resources remain authoritative, they are often difficult for the general public to understand. In contrast, short-video platforms like TikTok and Bilibili have emerged as key channels for disseminating health information due to their simplicity, vivid visuals, and interactivity⁸. By presenting complex health issues in an easily understandable format and providing quick access to information, short videos cater to the public's demand for convenience and immediacy in health-related content⁹. However, alongside the benefits of these platforms in spreading health information, users also face challenges regarding the quality and reliability of the content¹⁰. Due to the lack of peer review and stringent regulatory mechanisms, many

videos are of varying quality and may even propagate misleading or false information, which could mislead viewers' understanding of diseases, treatments, and health practices¹¹. This poses a significant risk of misinforming the public and negatively influencing health behaviors. Despite the growing role of short-video platforms in health communication, the quality and reliability of content related to medical topics, such as POF, remain underexplored.

This study aims to fill this gap by conducting a comprehensive evaluation of POF-related videos on TikTok and Bilibili. The objective is to analyze the scientific accuracy, reliability, and potential impact of these videos on public health education. By providing scientific evidence and a systematic analysis, this study seeks to contribute to improving the quality of health information on short-video platforms and support the development of policies and regulatory frameworks for platform governance.

Materials and Methods

Search Strategy and Data Extraction

The primary goal of this study was to gather short video data on the topic of "Premature Ovarian Failure" from Bilibili (<https://www.bilibili.com>) and TikTok (<https://www.douyin.com>). To ensure the comprehensiveness and representativeness of the data, the Chinese term for "Premature Ovarian Failure" was used as the search keyword. Data collection was carried out separately on both platforms. To minimize bias from personalized recommendations, all searches were conducted under newly created accounts, and searches were performed in an incognito mode, ensuring that no historical data or personalized algorithms influenced the results. Initially, videos ranked in the top 100 on each platform were collected based on the default sorting order. Following this, videos were excluded according to the following predefined criteria: 1) advertisements; 2) content irrelevant to the research topic; 3) videos uploaded within the last week, to maintain the stability and timeliness of the data. After applying these criteria, the final dataset of videos meeting the study's requirements was determined. During data extraction, the following variables were recorded: video source (Bilibili or TikTok), upload time, video duration (in seconds), number of likes, comments, shares, and collections (S1 Appendix). All data were extracted using the public APIs provided by the platforms, ensuring accuracy and consistency. These variables provided strong support for subsequent content analysis.

Video Quality Assessment

To evaluate the reliability and quality of the collected short videos, this study employed the modified DISCERN and Global Quality Score (GQS) assessments. Using the modified DISCERN, evaluators assessed whether each video met the following criteria: clarity, relevance, traceability, robustness, and fairness. Each criterion was rated as "yes" (1 point) or "no" (0 points), with a cumulative score calculated (ranging from 0 to 5 points)¹². The GQS utilized a 5-point Likert scale to evaluate the

overall quality of the video, with scores ranging from 1 (very poor) to 5 (very good). This scale considered aspects such as the professionalism of the video, the comprehensiveness of the information, the clarity of the presentation, and the viewer's ability to understand the content¹³. Furthermore, the completeness of the videos was assessed based on whether they included the following information: epidemiology, etiology, symptoms, diagnosis, treatment, and prevention (S2 Appendix). Videos were also evaluated based on how thoroughly the uploader explained the relevant topics, with three possible ratings: no explanation (0 points), partial explanation (1 point), and full explanation (2 points). Specialists, including gynecologists and reproductive medicine doctors, were distinguished from non-specialists, who were mainly from other medical fields or traditional Chinese medicine. All evaluations were conducted by two assessors with relevant medical backgrounds, who underwent unified training prior to scoring to ensure consistency in evaluation standards and minimize bias.

Statistical Analysis

This study employed descriptive statistics and nonparametric tests to analyze video characteristics and quality metrics. Categorical variables were presented as frequencies and percentages, whereas continuous variables were described using medians with interquartile ranges (IQRs). Chi-square tests or Fisher's exact tests were utilized to assess differences in platform distributions, while Mann-Whitney U tests compared interaction data and quality scores between TikTok and Bilibili. For multi-group comparisons among expert, non-expert, and individual user videos, Kruskal-Wallis tests were applied, followed by post-hoc Dunn's tests for pairwise comparisons. Cohen's kappa coefficients evaluated inter-rater reliability for GQS and modified DISCERN scores, with kappa values ≥ 0.8 indicating excellent agreement. Spearman correlation analysis explored relationships between GQS and modified DISCERN scores. Statistical computations were conducted using IBM SPSS Statistics 27.0, while GraphPad Prism 10.4 generated visualizations. A significance threshold of $P < 0.05$ was adopted, with no Bonferroni correction applied to post-hoc comparisons to preserve exploratory power. This analytical framework integrated robust nonparametric methods with advanced statistical software, ensuring alignment with data distributions and clinical research standards.

Results

Video Characteristics

Based on the established inclusion and exclusion criteria, the top 100 videos from TikTok and Bilibili were screened, resulting in a final sample of 187 videos. The detailed selection process is illustrated in Figure 1, while Table 1 presents the characteristics of the included videos. In terms of platform distribution, TikTok contributed 97 videos, while Bilibili provided 90. As for the creators, expert-uploaded videos were dominant, accounting for 61%, followed by non-expert videos at 19.8%, and

personal user videos at 19.3% (Figure 2). Overall, the videos showed modest engagement metrics. The median number of likes was 73 (11–224), the median number of comments was 11 (1–56), the median number of collections was 22 (6–89), and the median number of shares was 22 (3–52). These videos were generally short, with a median duration of 67 seconds (43–136). In terms of video quality, the median GQS score was 2 (1–3), and the median modified DISCERN score was 2 (1–3). As for inter-rater consistency, Cohen’s Kappa values for GQS and modified DISCERN scores were 0.916 and 0.938, respectively, indicating excellent agreement between the two evaluators and ensuring the reliability of the results.

Table1 Video characteristics.

Characteristic	N=187	
Short-video sharing platforms [n (%)]		
	Tiktok	97(51.9)
	Bilibili	90(48.1)
Video source [n (%)]		
	Specialists	114(61)
	Non-specialists	37(19.8)
	Individual user	36(19.3)
Number of likes [median (IQR)]		73(11-224)
Number of comments [median (IQR)]		11(1-56)
Number of collections [median (IQR)]		22(6-89)
Number of shares [median (IQR)]		22(3-52)
Video duration [s, median (IQR)]		67(43-136)
GQS scores [median (IQR)]		2(1-3)
modified DISCERN scores [median (IQR)]		2(1-3)

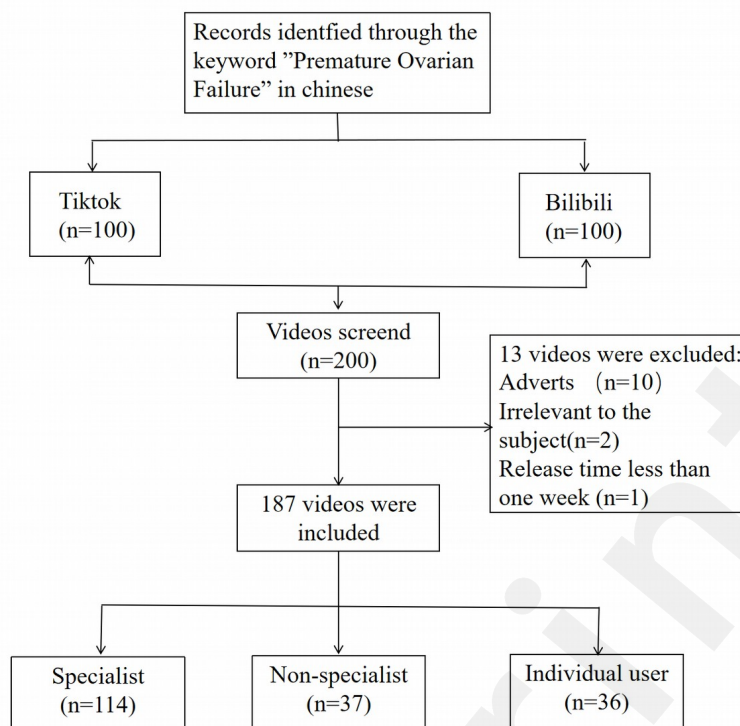


Fig1 The flow

chart of this study.

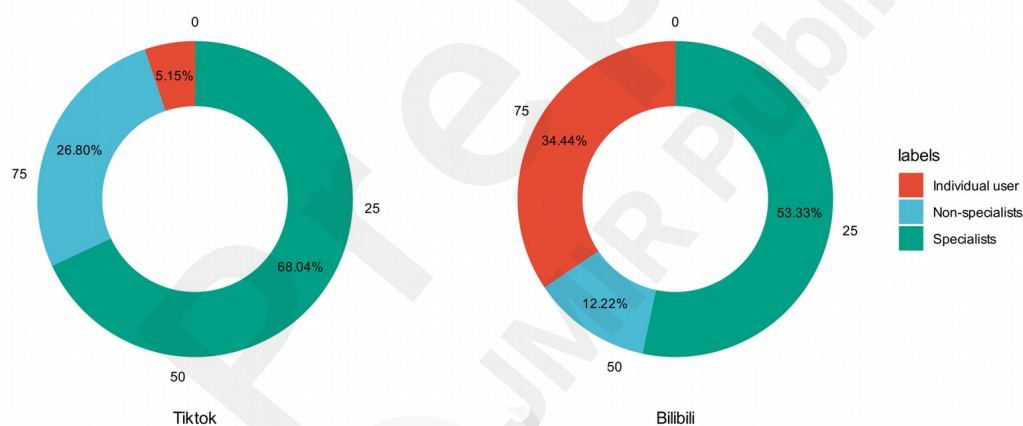


Fig2 Distribution

of videos by source type on TikTok and Bilibili platforms.

Video Content

Further analysis of the specific content covered in the selected short videos revealed that symptoms of POF were the most frequently discussed topic. Among all the videos, 29% provided a comprehensive and detailed explanation of POF symptoms. In addition to symptoms, videos focusing on diagnosis and treatment also garnered attention, with 16% and 7.4%, respectively, providing in-depth discussions on these aspects. In contrast, detailed information about the epidemiology and etiology of POF was rarely featured, with only a small number of videos briefly

mentioning these topics (Table 2).

Table2 Completeness of video content.

Video contents	Not involve (0 points)	Partial explanation (1 point)	Full explanation (2 points)
Epidemiology, n (%)	171□91.9□	14□7.4□	1□0.5□
Etiology, n (%)	135□72.1□	33□17.6□	19□10.1□
Symptoms, n (%)	57□30.4□	101□54.0□	29□15.5□
Diagnosis, n (%)	74□39.5□	83□44.3□	30□16.0□
Treatment, n (%)	93□49.7□	80□42.7□	14□7.4□
Prevention, n (%)	121□64.7□	54□28.8□	12□6.4□

Comparison of Features Across Platforms

Table 3 provides a detailed comparison of video features and quality between TikTok and Bilibili. The two platforms showed clear differences. On Bilibili, personal user-uploaded videos accounted for 34.4%, while on TikTok, this figure was only 5.2%. In terms of engagement, TikTok videos demonstrated significantly higher audience interaction than those on Bilibili. Specifically, the median number of likes, comments, collections, and shares on TikTok were 137 (62–331), 36 (10–79), 34 (11–115), and 31 (13–75), respectively, compared to Bilibili's median values of 12 (2–98), 2 (1–11), 15 (2–78), and 7 (1–32). Regarding video quality, TikTok videos had significantly higher GQS scores compared to Bilibili videos, although there were no statistical differences in the modified DISCERN scores between the two platforms (Figure 3).

Table3 Comparison of characteristics between different short-video platforms.

Variables	TikTok (N = 97)	Bilibili (N = 90)	p valve
Video source [n (%)]			<0.001
Specialists	66(68)	48(53.3)	
Non-specialists	26(26.8)	11(12.2)	
Individual user	5(5.2)	31(34.4)	
Number of likes [median (IQR)]	137□62-331□	12(2-98)	<0.001
Number of comments [median (IQR)]	36(10-79)	2(1-11)	<0.001
Number of collections [median (IQR)]	34(11-115)	15(2-78)	0.001
Number of shares [median (IQR)]	31(13-75)	7(1-32)	<0.001
Video duration [s, median (IQR)]	53(37-73)	122(63-244)	<0.001
GQS scores [median (IQR)]	3(2-3)	2(1-3)	0.004
modified DISCERN scores [median (IQR)]	2(1-3)	2(1-3)	0.317

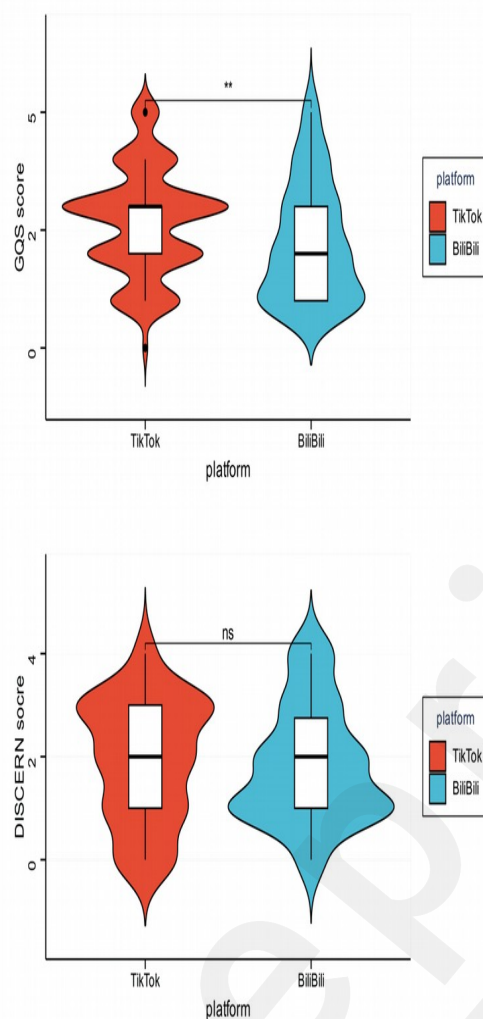


Fig3 Comparison of video quality scores between TikTok and Bilibili platforms.

Comparison of Features Across Different Video Sources

We further compared the features and quality of videos based on their source. In terms of video popularity, no statistically significant differences were observed in the number of likes, comments, collections, or shares across expert, non-expert, and personal user-uploaded videos (Table 4). However, significant differences were found in video length and quality. Expert-uploaded videos performed notably better in quality, with GQS and modified DISCERN scores of 3 (2–4) and 2 (1–3), respectively. Expert-uploaded videos had significantly higher GQS scores compared to non-expert and personal user-uploaded videos, with no significant differences between non-expert and personal user videos. As for modified DISCERN scores, expert-uploaded videos scored significantly higher than non-expert videos, though no significant difference was observed between expert and personal user-uploaded videos (Figure 4).

Table4 Comparison of different video source.

	Specialists (n=114)	Non-specialists (n=37)	Individual user (n=36)	p valve
--	------------------------	---------------------------	---------------------------	---------

Video duration [s, median (IQR)]	64(45-113)	54(37-85)	190(64-393)	<0.001
Number of likes [median (IQR)]	69(14-202)	97(29-196)	48(4-817)	0.659
Number of collections [median (IQR)]	23(5-81)	21(8-74)	27(8-857)	0.334
Number of comments [median (IQR)]	10(1-56)	19(4-47)	10(0-81)	0.610
Number of shares [median (IQR)]	22(2-48)	22(9-51)	21(2-178)	0.793
GQS scores [median (IQR)]	3(2-4)	2(1-2)	2(1-3)	<0.001
modified DISCERN scores [median (IQR)]	2(1-3)	1(0-2)	2(1-2)	<0.001

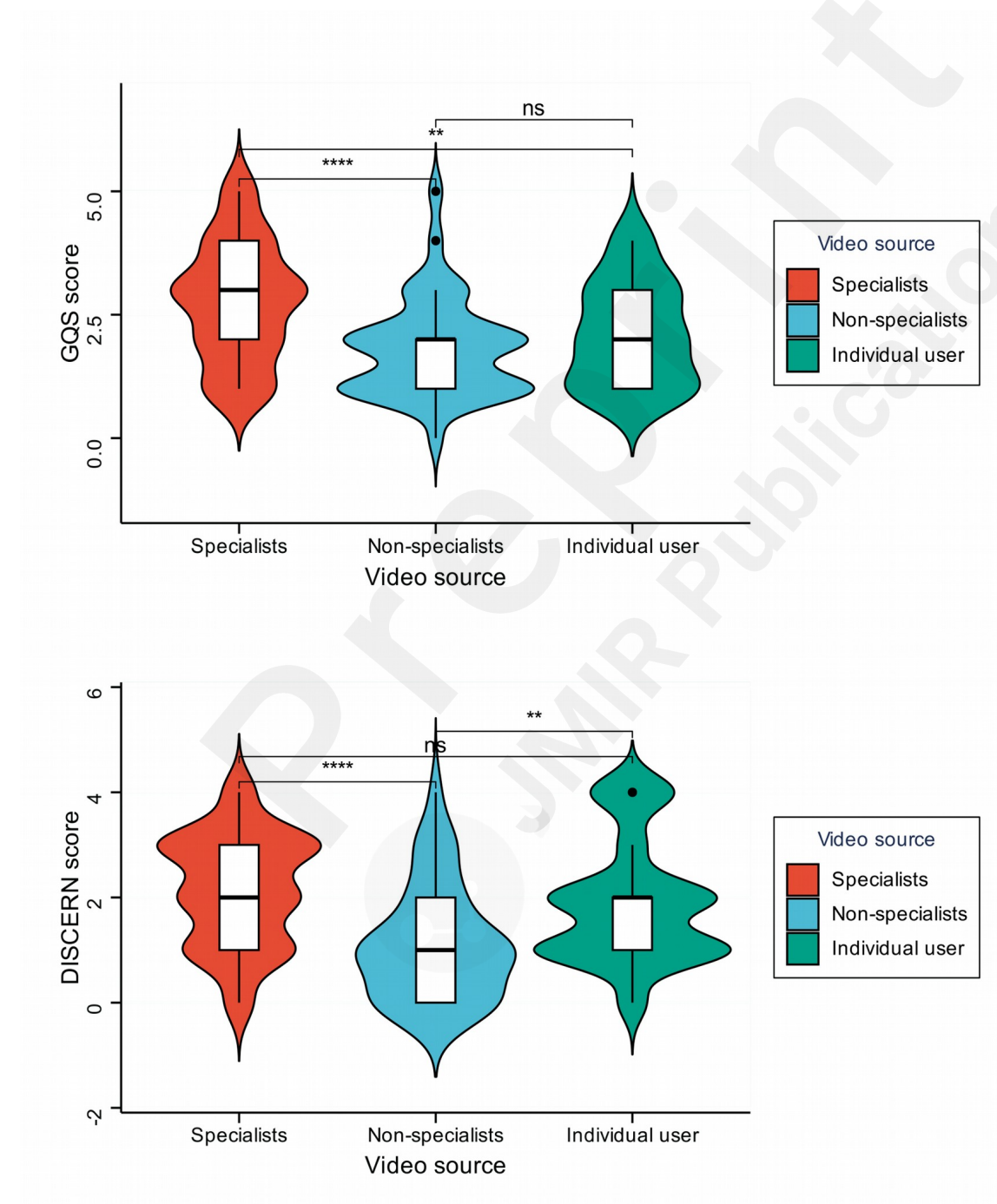


Fig4 Comparison of video quality scores by video source.

Correlation Analysis Between Video Features and Quality

Figure 5 presents the correlation analysis between video features and video quality across both TikTok and Bilibili platforms. On both platforms, a significant correlation was found between GQS and the modified DISCERN scores ($p < 0.001$). On TikTok, there was no notable correlation between video duration, likes, comments, collections, or shares with either GQS or the modified DISCERN scores. In contrast, on Bilibili, video duration was positively correlated with both GQS and modified DISCERN scores ($p < 0.001$), while no significant correlation was observed between likes, comments, collections, or shares with these quality measures.

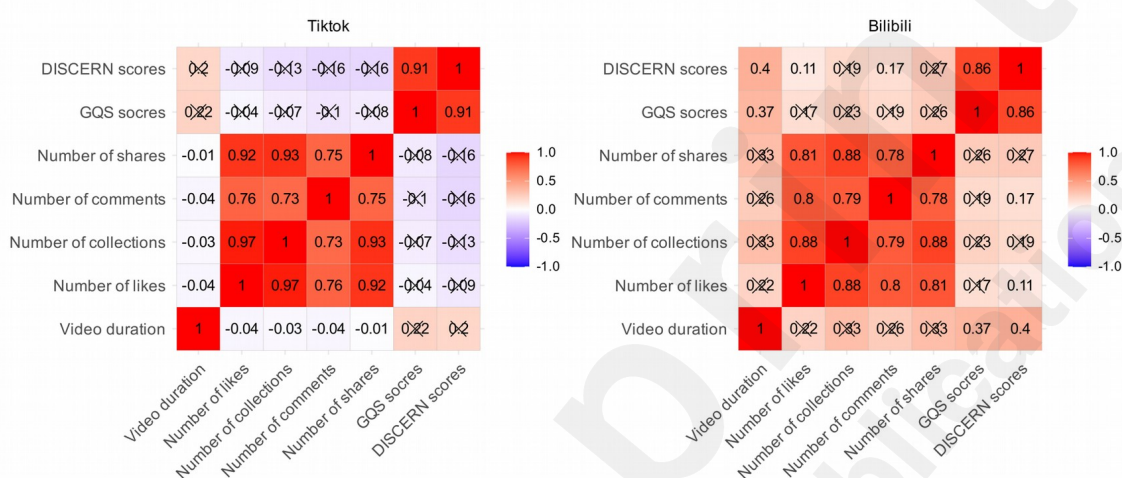


Fig5 Correlation matrix of video engagement metrics and quality scores on TikTok and Bilibili.

Discussion

With the rise of social media, especially short-video platforms, the way the public accesses health information has undergone a significant transformation^{14, 15}. Platforms like TikTok and Bilibili have become major channels for the dissemination of health-related content, attracting a large user base^{15, 16}. However, the quality and reliability of health videos on these platforms often vary, which can substantially influence the public's understanding of health matters¹⁷. While short-video platforms, through their visual and interactive content, have significantly improved the efficiency of information dissemination, they have also increased the risk of spreading misleading health information^{18, 19}. Therefore, evaluating the scientific accuracy, reliability, and audience interaction of health information shared on short-video platforms is not only clinically important but also holds substantial value for the development of health communication strategies and content moderation on these platforms²⁰.

This study conducted a comprehensive analysis of 187 health-related videos from TikTok and Bilibili, revealing notable differences in video quality, engagement data, and platform characteristics. First, the study found that TikTok videos had significantly higher median numbers of likes,

comments, collections, and shares compared to Bilibili, reflecting the differences in user engagement and interactivity across the two platforms. Furthermore, videos uploaded by experts consistently demonstrated higher quality. The health videos produced by experts scored highly on the GQS and modified DISCERN scales, indicating that the professional background of the uploader directly influences the content's quality and credibility. This finding aligns with previous studies, suggesting that expert-authored content tends to be more scientific and standardized²¹⁻²³. The use of GQS and modified DISCERN as evaluation tools showed that the overall quality of TikTok videos was superior to that of Bilibili videos, especially in terms of GQS scores. However, no significant differences were found in modified DISCERN scores between the two platforms, suggesting that while TikTok videos excelled in visual presentation and user interaction, their scientific accuracy and reliability were comparable to those found on Bilibili. This highlights the multidimensional nature of video quality: on one hand, the visual appeal and interactivity of a video are critical for attracting viewers, but on the other hand, its scientific validity and accuracy remain equally important^{24, 25}. Thus, both content quality and presentation should be prioritized when disseminating health information via short videos²⁶.

Through Spearman correlation analysis, this study further explored the relationship between video quality and user interaction data, such as likes, comments, collections, and shares. On TikTok, no significant correlations were observed between video duration, likes, comments, collections, shares, and video quality (GQS and modified DISCERN scores). This suggests that, although TikTok videos show high user interaction, these interactions do not directly contribute to enhancing the quality or scientific accuracy of the video content. Conversely, on Bilibili, video duration was positively correlated with both GQS and modified DISCERN scores. This finding indicates that longer videos may offer more valuable health information, thereby improving the overall quality score of the video^{10, 27}.

Several concerns regarding the scientific accuracy and misleading nature of the content about POF on short-video platforms were also identified. For instance, some videos exaggerated the effectiveness of certain treatments, such as herbal medicine and massage, which lack scientific evidence, potentially leading viewers to abandon conventional medical interventions. Additionally, some videos conflated the concepts of POF and Primary Ovarian Insufficiency (POI), which could mislead viewers about the differences between these conditions and affect their coping strategies²⁸. Furthermore, some videos provided overly simplistic lifestyle or dietary advice, neglecting individual differences, the multifactorial nature of diseases, and the importance of medical interventions. Lastly, a number of videos discussed symptoms of POF, such as irregular menstruation, without providing a comprehensive diagnostic process, which might lead viewers to misinterpret general physiological changes as POF, inducing unnecessary anxiety and panic. These misleading

contents not only affect the public's understanding of POF but may also lead to misguided health decisions and treatment delays. Therefore, ensuring the scientific accuracy and evidence-based nature of health information disseminated on short-video platforms is critical^{29, 30}.

The main significance of this study lies in its provision of a comprehensive method for evaluating the quality of health information on short-video platforms. By comparing TikTok and Bilibili, we have revealed how platform characteristics, video sources, and interaction data influence the quality of health information dissemination. Furthermore, by integrating the GQS and modified DISCERN tools, this study offers a more systematic and objective evaluation of health information videos, providing valuable insights for public health policy-making, content regulation on short-video platforms, and health education initiatives^{23, 31}. The results suggest that significant differences exist in user interaction and video quality across platforms, and that the professional background of video uploaders plays a decisive role in determining content quality³². These findings provide empirical support for optimizing health information dissemination strategies.

While this study offers valuable empirical data on the quality of health information on short-video platforms, it does have some limitations. First, this study focused only on TikTok and Bilibili, and did not encompass other short-video platforms, potentially limiting the scope of health information dissemination analysis^{33, 34}. Second, as short-video content is constantly being updated, the data used in this study may not fully reflect the latest trends and developments. Third, video quality ratings are somewhat subjective, and although consistency tests were conducted to ensure inter-rater reliability, the application of the rating tools may still be influenced by the evaluators' personal interpretations^{35, 36}. Additionally, this study primarily analyzed videos in Chinese, and future research could consider cross-linguistic and cross-cultural comparisons to explore how health information is disseminated in different cultural contexts. Lastly, the integration of AI-driven analysis tools could provide more objective and efficient solutions for video content analysis, a direction that future research may wish to explore.

Conclusion

This study assessed the quality of 187 videos related to ovarian insufficiency on TikTok and Bilibili using the GQS and modified DISCERN scoring systems. Overall, the videos on both platforms were found to be unsatisfactory in terms of quality and reliability. Notably, videos on TikTok were slightly better than those on Bilibili. Videos shared by specialists demonstrated higher quality and reliability, offering viewers more valuable information. Given the rise of short videos, it is crucial for healthcare professionals and institutions to ensure the availability of high-quality content related to ovarian insufficiency. Furthermore, short video platforms should strengthen their monitoring and review mechanisms. Patients should approach videos on Bilibili and TikTok with caution when seeking medical information.

Acknowledgements

The authors would like to express their gratitude to the participants who participated in the study.

Author contributions statement

Ren Xu did conceptualization and writing the original draft. Jianzhi Su did conceptualization, formal analysis, and validation. Yanan Ren performed data curation and formal analysis. Luyang Su data curation. Xinjun L was responsible for the software and visualization. All authors contributed to the article and approved the submitted version.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Statement

The data used in this study were sourced from publicly available video content published on platforms such as Bilibili and TikTok. These videos are publicly accessible, and no personal privacy information was involved during the data collection process. All analyzed content was publicly available, and the study did not involve the collection or processing of users' private information. In accordance with relevant ethical review guidelines, ethical approval for this study was not required.

Funding

The study is supported by Medical Science Research Project of Hebei (NO. 20240273).

Abbreviations

POF: Premature Ovarian Failure

GQS: Global Quality Score

References

1. Chen J, Zhou Q, Zhang Y, et al. Discovery of novel serum metabolic biomarkers in patients with polycystic ovarian syndrome and premature ovarian failure. *Bioengineered* 2021; 12: 8778-8792. DOI: 10.1080/21655979.2021.1982312.
2. Chen M, He QD, Guo JJ, et al. Electro-Acupuncture Regulates Metabolic Disorders of the Liver and Kidney in Premature Ovarian Failure Mice. *Front Endocrinol (Lausanne)* 2022; 13: 882214. 20220725. DOI: 10.3389/fendo.2022.882214.
3. Wang X, Wang L and Xiang W. Mechanisms of ovarian aging in women: a review. *J Ovarian Res* 2023; 16: 67. 20230406. DOI: 10.1186/s13048-023-01151-z.
4. Zhang M, Yu X, Li D, et al. Nrf2 Signaling Pathway Mediates the Protective Effects of Daphnetin Against D-Galactose Induced-Premature Ovarian Failure. *Front Pharmacol* 2022; 13: 810524. 20220128. DOI: 10.3389/fphar.2022.810524.
5. Ghahremani-Nasab M, Ghanbari E, Jahanbani Y, et al. Premature ovarian failure and tissue engineering. *J Cell Physiol* 2020; 235: 4217-4226. 20191029. DOI: 10.1002/jcp.29376.
6. He Z, Wang Z, Song Y, et al. The Reliability and Quality of Short Videos as a Source of Dietary Guidance for Inflammatory Bowel Disease: Cross-sectional Study. *J Med Internet Res* 2023; 25: e41518. 20230209. DOI: 10.2196/41518.
7. Mohamed F and Shoufan A. Users' experience with health-related content on YouTube: an exploratory study. *BMC Public Health* 2024; 24: 86. 20240103. DOI: 10.1186/s12889-023-17585-5.
8. Liu H, Peng J, Li L, et al. Assessment of the reliability and quality of breast cancer related videos on TikTok and Bilibili: cross-sectional study in China. *Front Public Health* 2023; 11: 1296386. 20240122. DOI: 10.3389/fpubh.2023.1296386.
9. Sun F, Zheng S and Wu J. Quality of Information in Gallstone Disease Videos on TikTok: Cross-sectional Study. *J Med Internet Res* 2023; 25: e39162. 20230208. DOI: 10.2196/39162.
10. Chen J and Wang Y. Social Media Use for Health Purposes: Systematic Review. *J Med Internet Res* 2021; 23: e17917. 20210512. DOI: 10.2196/17917.

11. Memioglu T and Ozyasar M. Analysis of YouTube videos as a source of information for myocarditis during the COVID-19 pandemic. *Clin Res Cardiol* 2022; 111: 1113-1120. 20220426. DOI: 10.1007/s00392-022-02026-x.
12. Kanzow P, Büttcher AF, Wiegand A, et al. Quality of Information Regarding Repair Restorations on Dentist Websites: Systematic Search and Analysis. *J Med Internet Res* 2020; 22: e17250. 20200415. DOI: 10.2196/17250.
13. Ma M, Yin S, Zhu M, et al. Evaluation of Medical Information on Male Sexual Dysfunction on Baidu Encyclopedia and Wikipedia: Comparative Study. *J Med Internet Res* 2022; 24: e37339. 20220809. DOI: 10.2196/37339.
14. Liu Q, Zheng Z, Zheng J, et al. Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach. *J Med Internet Res* 2020; 22: e19118. 20200428. DOI: 10.2196/19118.
15. Yuan Y and Wang Q. Characteristics, hotspots, and prospects of short video research: A review of papers published in China from 2012 to 2022. *Heliyon* 2024; 10: e24885. 20240124. DOI: 10.1016/j.heliyon.2024.e24885.
16. Bai G, Fu K, Fu W, et al. Quality of Internet Videos Related to Pediatric Urology in Mainland China: A Cross-Sectional Study. *Front Public Health* 2022; 10: 924748. 20220615. DOI: 10.3389/fpubh.2022.924748.
17. Juvalta S, Kerry MJ, Jaks R, et al. Electronic Health Literacy in Swiss-German Parents: Cross-Sectional Study of eHealth Literacy Scale Unidimensionality. *J Med Internet Res* 2020; 22: e14492. 20200313. DOI: 10.2196/14492.
18. Liang J, Wang L, Song S, et al. Quality and Audience Engagement of Takotsubo Syndrome-Related Videos on TikTok: Content Analysis. *J Med Internet Res* 2022; 24: e39360. 20220926. DOI: 10.2196/39360.
19. Qi Y, Han J, Lu X, et al. A study on satisfaction evaluation of Chinese mainstream short video platforms based on grounded theory and CRITIC-VIKOR. *Heliyon* 2024; 10: e30050. 20240424. DOI: 10.1016/j.heliyon.2024.e30050.
20. Chen Y, Wang Q, Huang X, et al. The quality and reliability of short videos about thyroid nodules on Bilibili and TikTok: Cross-sectional study. *Digit Health* 2024; 10: 20552076241288831. 20241007. DOI: 10.1177/20552076241288831.
21. Shi A, El Haddad J, Cai P, et al. Mpox (monkeypox) information on TikTok: analysis of quality and audience engagement. *BMJ Glob Health* 2023; 8. DOI: 10.1136/bmjgh-2022-011138.
22. Matthews MR, Abdulbaki H, Ryan WR, et al. Oropharyngeal Cancer and the HPV Vaccine: Analysis of Social Media Content. *Laryngoscope* 2025 20250217. DOI: 10.1002/lary.32076.
23. Song S, Xue X, Zhao YC, et al. Short-Video Apps as a Health Information Source for Chronic Obstructive Pulmonary Disease: Information Quality Assessment of TikTok Videos. *J Med Internet Res* 2021; 23: e28318. 20211220. DOI: 10.2196/28318.
24. Ayers SL and Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. *Health (London)* 2007; 11: 327-347. DOI: 10.1177/1363459307077547.
25. Chen YY, Li CM, Liang JC, et al. Health Information Obtained From the Internet and Changes in Medical Decision Making: Questionnaire Development and Cross-Sectional Survey. *J Med Internet Res* 2018; 20: e47. 20180212. DOI: 10.2196/jmir.9370.
26. Elhariry M, Malhotra K, Goyal K, et al. A SIMBA CoMICs Initiative to Cocreating and Disseminating Evidence-Based, Peer-Reviewed Short Videos on Social Media: Mixed Methods Prospective Study. *JMIR Med Educ* 2024; 10: e52924. 20241030. DOI: 10.2196/52924.
27. Yeung AWK, Tosevska A, Klager E, et al. Medical and Health-Related Misinformation on Social Media: Bibliometric Study of the Scientific Literature. *J Med Internet Res* 2022; 24: e28152. 20220125. DOI: 10.2196/28152.
28. Wang Y, Jiang J, Zhang J, et al. Research Progress on the Etiology and Treatment of Premature Ovarian Insufficiency. *Biomed Hub* 2023; 8: 97-107. 20231213. DOI: 10.1159/000535508.
29. Aktas BK, Demirel D, Celikkaleli F, et al. YouTube™ as a source of information on prostatitis: a quality and reliability analysis. *Int J Impot Res* 2024; 36: 242-247. 20230116. DOI: 10.1038/s41443-023-00666-9.
30. Househ M, Borycki E and Kushniruk A. Empowering patients through social media: the benefits and challenges. *Health Informatics J* 2014; 20: 50-58. DOI: 10.1177/1460458213476969.
31. Alhlal AS and Alzghaibi HA. Evaluating the effectiveness and limitations of online health information tools in assessing the quality of medication-related content. *Front Public Health* 2025; 13: 1460202. 20250224. DOI: 10.3389/fpubh.2025.1460202.
32. Enver N, Doruk C, Kara H, et al. YouTube™ as an information source for larynx cancer: a systematic review of video content. *Eur Arch Otorhinolaryngol* 2020; 277: 2061-2069. 20200316. DOI: 10.1007/s00405-020-05906-y.
33. 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. DOI: 10.1007/s11695-019-03738-2.
34. Prestin A, Vieux SN and Chou WY. Is Online Health Activity Alive and Well or Flatlining? Findings From 10 Years of the Health Information National Trends Survey. *J Health Commun* 2015; 20: 790-798. 20150604. DOI: 10.1080/10810730.2015.1018590.
35. Mueller SM, Hongler VNS, Jungo P, et al. Fiction, Falsehoods, and Few Facts: Cross-Sectional Study on the Content-Related Quality of Atopic Eczema-Related Videos on YouTube. *J Med Internet Res* 2020; 22: e15599. 20200424. DOI: 10.2196/15599.
36. Guan JL, Xia SH, Zhao K, et al. Videos in Short-Video Sharing Platforms as Sources of Information on Colorectal Polyps: Cross-Sectional Content Analysis Study. *J Med Internet Res* 2024; 26: e51655. 20241029. DOI: 10.2196/51655.

Preprint
JMIR Publications

Supplementary Files