scientific reports

Check for updates

OPEN

Evaluating the quality of TikTok videos on coronary artery disease using various scales to examine correlations with video characteristics and high-quality content

Peng Hao[1,2], Guixin Liu[3], Shuang Lian[4], Jiaxu Huang[5] & Lin Zhao[1,2]✉

Background Coronary artery disease (CAD) is a major public health concern, yet reliable sources of relevant information are limited. TikTok, a popular social media platform in China, hosts diverse health-related videos, including those on CAD; however, their quality varies and is largely unassessed. **Objective** This study aimed to investigate the quality of CAD-related videos on TikTok and explore the correlation between video characteristics and high-quality videos. **Methods** A total of 122 CAD-related short videos on TikTok were analyzed on July 18, 2023. Basic video information and sources were extracted. Two evaluators independently scored each video using DISCERN (a health information quality scale), the Patient Education Materials Assessment Tool (PEMAT) and the Health on the Net (HONcode) scales. Videos were categorized into four groups based on their source, with the medical professional group further categorized by job titles. Simple linear analysis was used to examine the linear relationship across different scales and to explore the relationship between video characteristics (video length, time since posting, the number of "likes", comments and "favorites", and the number of followers of the video creator) and different scales. **Results** AQVideos were categorized into four groups based on their source: medical professionals (n = 98, 80.3%), user-generated content (n = 11, 9.0%), news programs (n = 4, 3.3%), and health agencies or organizations (n = 9, 7.4%). The score of DISCERN was 46.5 ± 7.6/80, the score rate of PEMAT was 79.2 ± 12.6%/100%, and the number of score items for HONcode was 1.4 ± 0.6/8. In Sect. 1 of DISCERN, user-generated content scored highest (29.1 ± 3.6), followed by medical professionals (28.6 ± 2.4), health agencies or organizations (28.0 ± 0.0) and news programs (28.0 ± 0.0)($P = 0.047$). In HONcode, most videos met only one or two of the eight evaluation criteria. PEMAT scores varied slightly across categories without significant differences ($P = 0.758$). Medical professionals were further divided into senior (n = 69, 70.4%) and intermediate (n = 29, 29.6%) groups, with intermediate professionals scoring higher in DISCERN ($P < 0.001$). In simple linear analysis models, no linear correlation was found between DISCERN and PEMAT scores ($P = 0.052$). Time since posting on TikTok was negatively correlated with DISCERN ($P = 0.021$) and PEMAT scores ($P = 0.037$), and the number of "favorites" was positively correlated to DISCERN score ($P = 0.007$). **Conclusion** The quality of CAD-related videos on China's TikTok is inconsistent and varies across different evaluation scales. Videos posted by medical professionals with intermediate titles tended to offer higher quality, more up-to-date content, as reflected by higher "favorite" counts. HONcode may not be suitable for short video evaluation due to its low score rate, while DISCERN and PEMAT may be effective tools for short video evaluation. However, their lack of consistency in evaluation dimensions highlight the need for a tailored scoring system for short videos.

Keywords TikTok, Short video, Quality evaluation, DISCERN, PEMAT, HONcode

Abbreviations

APCS	Acute Pancreatitis Content Score
CAD	Coronary artery disease

CRAAP	Currency, Relevance, Authority, Accuracy, and Purpose
HONcode	Health on the Net Code
PEMAT	Patient Education Materials Assessment Tool
SD	standard deviation

[1]Department of Cardiology, Beijing Chaoyang Hospital, Capital Medical University, Beijing 100020, China. [2]Heart Center and Beijing Key Laboratory of Hypertension, Beijing Chaoyang Hospital, Capital Medical University, Beijing 100020, China. [3]Beijing Anzhen Hospital, Capital Medical University, Beijing 100029, China. [4]Geriatric Center, Affiliated Hospital of Inner Mongolia Medical University, Hohhot City 010059, Inner Mongolia Autonomous Region, China. [5]Mudan People's Hospital, Heze City 274000, Shandong Province, China. ✉email: trichina2007@126.com

Coronary artery disease (CAD) is a major public health concern worldwide and is the leading cause of morbidity and mortality in many countries[1]. In China, there has been a tremendous increase in the incidence of CAD, due to aging demographics, unhealthy lifestyles, and environmental changes after several decades of rapid economic development[2]. According to the Report on Cardiovascular Health and Diseases in China 2023, the age-standardized incidence rate of CAD increased from 177.1 per 100,000 person-years in 1990 to 203.7 per 100,000 person-years in 2010, before slightly declining to 197.4 per 100,000 person-years in 2019 [https://doi.org/10.3967/bes2024.162]. According to the China Health Statistics Yearbook 2020[3], the mortality rates of CAD were 121.59/100,000 in urban and 130.14/100,000 in rural areas in 2019. CAD is associated with substantial economic impact, including treatment costs and productivity loss due to illness or premature mortality[4]. In China, the high cost of healthcare and the limited insurance coverage pose challenges in accessing CAD treatment[5]. Meanwhile, the lack of reliable channels for disseminating CAD-related knowledge about the prevention, recognition, and treatment of CAD significantly contributes to the elevated risk in China[6,7]. To reduce the morbidity and mortality rates, it is crucial to improve public understanding of CAD.

The Chinese version of TikTok "抖音" (referred to as TikTok hereafter) is a popular short video platform in China that allows users to create and share videos through personalized profiles or pages, providing a widely accessible medium for disseminating health information[8,9]. Studies have shown that TikTok hosts popular science content on various diseases, including diabetes, chronic obstructive pulmonary disease, and anxiety[10–12]. However, the quality of such videos is inconsistent. Previous studies evaluating the quality of videos primarily used scales such as DISCERN (a health information quality scale), Patient Education Materials Assessment Tool (PEMAT), The Currency, Relevance, Authority, Accuracy, and Purpose (CRAAP), and the Health on the Net (HONcode). These tools assess video content from multiple perspectives, focusing on both the source and the content, with low scores denoting poor quality: DISCERN scores below 3 out of 5 indicate unreliable content and insufficient treatment information; PEMAT scores below 70% indicate poor patient understandability and actionability; CRAAP scores below 60% reflect issues with the timeliness and credibility of the information; and HONcode compliance below 4 out of 8 principles suggests a lack of trustworthiness of the information[13,14]. Collectively, these scales provide a comprehensive assessment of whether a video offer accurate, accessible, and actionable health information to viewers. Notably, previous research has found that most TikTok videos scored poorly on DISCERN and PEMAT[15], with some even being misleading[16]. This raises concerns about the quality of CAD-related videos on TikTok.

Previous studies have primarily focused on the quality evaluation of long videos[13,17,18], leaving it unclear whether the scales used to evaluate long videos can effectively assess the quality of short videos. This study aims to evaluate the quality of CAD-related short videos on TikTok. To objectively evaluate short videos based on different criteria, we selected three international video evaluation scales (DISCERN, PEMAT and HONcode) to assess the quality of CAD-related videos and explored the consistency between these scales. Considering the limited research on the relationship between video characteristics and the quality of short videos, this study also evaluated the correlations between video characteristics (video length, time since posting, the number of "likes", comments and "favorites", and the number of followers of the video creator) and different evaluation scales.

Methods

Search strategy and data extraction

We searched the keyword "冠心病" (the most commonly used term for CAD in Chinese medical communication) on TikTok on July 18, 2023, and retrieved 133 videos based on relevance ranking by TikTok's algorithm. Cookies and browsing history were cleared before each search to avoid algorithmic bias. The reviewers were blinded to video sources during evaluation. Both pinned comments and video captions were included in the evaluation.

Following established sample sizes from similar studies[10,11,15,19,20], 122 videos directly related to CAD were included for further data extraction and analysis. Irrelevant content ($n=3$), commercial advertisements ($n=4$), videos with no sound ($n=1$), and videos in languages other than Chinese ($n=3$) were excluded (Fig. 1). Basic information for each video was extracted, including video length, time since posting, the number of "likes", comments and "favorites", the number of followers of the video creator, name and type of the video creator (individual or organization), and the titles of medical professional video creators, who comprised the largest proportion of video creators.

The verification of video sources was conducted through a systematic process. For medical professionals, we verified their identity using two criteria: official verification badge ("V" mark) on their TikTok profiles, which requires submission of professional certification to the platform, and the work environment shown in their videos, such as hospital settings and clinical scenarios. For organizational accounts, verification was based on their official organizational verification status on TikTok and institutional business license information displayed in their profiles.

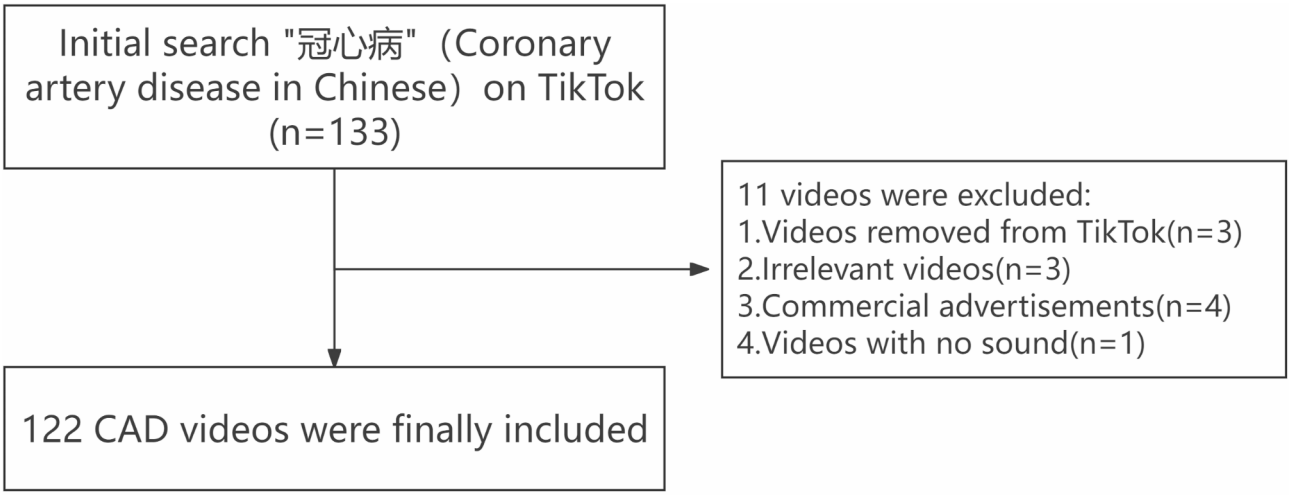

Fig. 1. Video screening flowchart.

Study groups

We measured two aspects of CAD-related videos on TikTok: their characteristics and information quality. Videos were categorized into four groups based on their source: medical professionals, user-generated content, news programs, and health agencies or organizations. Given the large proportion of medical professional creators, we further categorized them into senior and intermediate groups based on their professional titles, as determined according to China's professional skill classification system. The senior classification corresponds to an attending physician in the U.S. system, while the intermediate classification is comparable to a fellow physician in the U.S.

Quality assessments

Two evaluators assessed each video independently and scored each item on the scales to assess video quality using the DISCERN, PEMAT, and HONcode instruments. A comprehensive explanation for each sub-item of the rating scales is provided in Supplementary Materials.

DISCERN, established in 1998, is one of the most widely adopted tools for assessing the quality of health information, particularly regarding treatment choices. Comprising 16 questions across three sections, the instrument employs a scale from 1 (poor) to 5 (good)[21]. The initial eight questions pertain to the publication's reliability, assessing its trustworthiness as an information source for treatment decisions. The subsequent seven questions delve into treatment details, with scores reflecting information quality. The third section concludes with a single question, inviting users to rate the publication's overall quality.

The PEMAT is a novel tool designed for evaluating the comprehensibility and feasibility of patient education materials, including both print and audiovisual formats, covering a range of topics. Its development involved extensive input from a panel of experts in fields such as health literacy, health communication, content creation across different mediums, patient education, patient engagement, and health information technology. The instrument comprises 26 questions, with the first 19 addressing understandability. A higher score indicates greater clarity and comprehensibility. The subsequent seven questions in the second section assess the ease of implementing the treatments presented in the video materials[22]. In our study, we used the 17 video-applicated questions to evaluate the video by counting the scoring rate (%).

HONcode stands as the most time-tested and widely employed ethical framework, ensuring the credibility and integrity of medical and health-related content on the internet[23]. We employed eight adapted HONcode principles to evaluate the quality of the videos. These principles cover several crucial aspects, including the authority of the information source, the platform's purpose, confidentiality, references and their currency, fairness, contact information for the platform, funding disclosure, and transparent labeling of advertisements[24]. Each evaluator assessed the adherence of each video to these principles, using a binary scale ("yes" = 1; "no" = 0). Subsequently, we calculated a cumulative adherence score to quantify the total number of HONcode principles upheld by each video. Details of the scales can be found in Supplementary Materials.

Statistical analysis

Normally distributed variables were compared among the groups using ANOVA test. Non-normally distributed variables were compared among the groups using Kruskal-Wallis test. Variables in the characteristics of videos were presented as median. In each scale, scores were presented as mean ± standard deviation (SD). Simple linear regression analysis was performed to explore the association between DISCERN, PEMAT, and HONcode instruments. Due to the low score of the HONcode instruments (1.4/8), it may not be suitable for evaluating short videos. Therefore, only the results of the correlation between PEMAT and DISCERN are presented in the main text, while the regression analyses between HONcode and PEMAT/DISCERN can be found in Supplementary Materials.

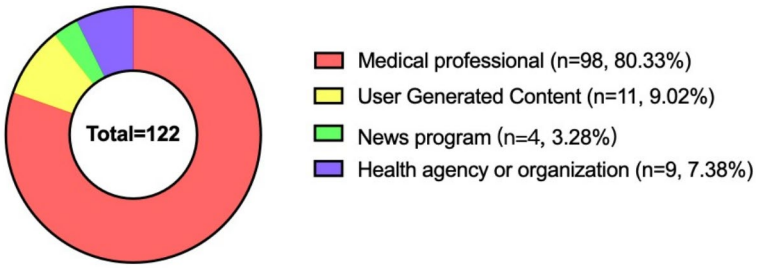

Fig. 2. Sources of reviewed CAD-related videos on TikTok.

	Medical professionals ($n=98$)	User-generated content ($n=11$)	News programs ($n=4$)	Health agencies or organizations ($n=9$)	Total ($n=122$)	*P* value
Video length (seconds), median	85.5	85.0	52.5	57.0	76.5	0.110
Time since posting (months), median	13.0	19.0	18.0	13.0	14.0	0.352
Number of "likes", median	1086.5	4414.0	2999.0	1954.0	1255.0	0.427
Number of comments, median	63.5	212.0	42.0	36.0	63.5	0.262
Number of "favorites", median	133.5	323.0	84.0	157.0	160.0	0.160
Number of followers of the video creator, median	443,000.0	559,000.0	1,223,500.0	750,000.0	471,500.0	0.787

Table 1. Characteristics of CAD-related TikTok videos by sources. CAD, coronary artery disease.

Simple linear regression analysis was also performed to assess the association between video characteristics (video length, time since posting, the number of "likes", comments and "favorites", and the number of followers of the video creator) and scales (DISCERN, PEMAT, and HONcode). The results were expressed by the equation and *P* value in simple linear regression model. $P<0.05$ was considered statistically significant.

To explore temporal trends in video quality, we categorized the videos into three time periods: 2023–present ($n=54$), 2022–2023 ($n=39$), and before 2022 ($n=29$). We compared the total scores of DISCERN, HONcode, and PEMAT across these periods to examine how content quality evolved over time.

Statistical analyses were performed using Stata version 15.0 software (4905 Lakeway Drive, College Station, Texas 77845 USA). The interrater reliability for each item of the three instrument ranges is more than 0.90, and all reliability coefficients are significant at an error margin of 0.1%. The interrater reliability is satisfied.

Results

Video sources

CAD-related videos on TikTok came from four types of sources: medical professionals, user-generated content, news programs, and health agencies or organizations. Medical professionals posted the largest proportion of videos ($n=98$, 80.3%), followed by user-generated content ($n=11$, 9.0%). Health agencies or organizations and news programs posted a total of 13 videos (10.7%), with health agencies or organizations posting the majority ($n=9$, 7.4%), followed by news programs ($n=4$, 3.3%) (Fig. 2).

Video characteristics

Table 1 shows the characteristics of the videos by source. We analyzed the following characteristics: video length, time since posting, the number of "likes", comments and "favorites", and the number of followers of the video creator. Video length ranged from 8 to 841 s, with an average length of 85.5 s. All videos were published after 2018, with the oldest video being on TikTok for 59 months and the most recent posted less than one month before data collection. The number of "likes" ranged from 0 to 879,000, comments from 0 to 37,000 and "favorites" from 0 to 62,000. The number of followers for each creator ranged from 68 to 10,473,000. No statistical differences were observed between groups for any video characteristics ($P>0.05$) (Table 1).

Information quality

Table 2 showed eight adapted HONcode principles and the frequency of adherence in the 122 reviewed videos. The mean number of score items for HONcode was 1.4 ± 0.6. All 122 videos (100%) met criteria #1, 39 (32%) met criteria #2, 6 (4.9%) met criteria #3, 2 (1.6%) met criteria #4, 2 (1.6%) met criteria #5, and no video met criteria #6–8 (Table 2).

The mean score of DISCERN was 46.5 ± 7.6 in all 122 videos. The mean scores of Sect. 1 (reliability of the videos) indicate significant differences across video sources ($P=0.047$). User-generated content had the highest scores (29.1 ± 3.6), followed by videos posted by medical professionals (28.6 ± 2.4), while the videos posted by health agencies or organizations and news programs had the lowest scores (both 28.0 ± 0.0). The mean scores of Sect. 2 (quality of treatment choices, items 9–15) and Sect. 3 (overall information quality, item 16) showed no significant statistical significance among different sources (all $P>0.05$). There was no significant difference in

Number	Criteria	Number of videos in compliance (% of total sample)
1	Any medical or health advice given in the video must come from a qualified health professional unless it is clearly stated the information does not come from a qualified health source.	122 (100.0)
2	The information provided in the videos must be designed to support the patient's CAD self-management but is not meant to replace a patient-physician relationship.	39 (32.0)
3	The information in the video maintains the right to confidentiality and respect of the individual patient featured.	6 (4.9)
4	Each video contains references to source data on information presented or contains a specific HTML link to source information.	2 (1.6)
5	Each video containing claims on the benefits or performance of specific, skills/behaviors, interventions, treatments, products, etc. must be supported by evidence through references or HTML links.	2 (1.6)
6	The video must provide the viewer with contact information, or a website link to more information.	0 (0.0)
7	Any individuals or organizations that contribute funds, services, or material in the posted video must be clearly identified in the video or video description.	0 (0.0)
8	If advertisement supports funding to the video or the video's developers, it must be clearly stated. Included advertising must be clearly differentiable to the viewer: there should be a clear difference between the advertising material and the educational material in the video.	0 (0.0)

Table 2. Eight adapted HONcode principles and frequency of adherence in reviewed CAD-related TikTok videos ($n = 122$). CAD, coronary artery disease; HONcode: Health on the Net Code; HTML, hypertext markup language; URL, uniform resource locator.

	Total ($n = 122$)	Medical professionals ($n = 98$)	User-generated content ($n = 11$)	News programs ($n = 4$)	Health agencies or organizations ($n = 9$)	*P* value
DISCERN						
Reliability of the videos (items 1–8), mean (SD)	28.6 (2.4)	28.6 (2.4)	29.1 (3.6)	28.0 (0.0)	28.0 (0.0)	0.047
Quality of treatment choices (items 9–15), mean (SD)	14.7 (6.5)	14.8 (6.4)	14.1 (8.9)	14.3 (6.2)	14.3 (5.7)	0.463
Overall information quality (item 16), mean (SD)	3.3 (0.8)	3.3 (0.8)	3.5 (0.8)	3.2 (1.0)	3.0 (0.8)	0.951
Total DISCERN score, mean (SD)	46.5 (7.6)	46.7 (7.6)	46.6 (10.2)	45.3 (5.7)	45.6 (6.8)	0.486
PEMAT						
Understandability (items 1, 3–5, 8–14, 18–19), mean (SD) [scoring rate (%)]	60.0 (9.6)	60.3 (9.6)	59.0 (10.2)	58.2 (8.2)	58.5 (11.5)	0.868
Actionability (items 20–22, 25), mean (SD) [scoring rate (%)]	19.2 (6.4)	19.5 (6.2)	16.8 (8.5)	21.4 (6.2)	18.2 (6.6)	0.586
Total PEMAT score, mean (SD) [scoring rate (%)]	79.2 (12.6)	79.8 (12.5)	75.8 (11.8)	79.6 (12.6)	76.8 (16.2)	0.758
HonCode						
Total HonCode score, mean (SD)	1.4 (0.6)	1.4 (0.6)	1.3 (0.5)	1.4 (0.5)	1.3 (0.5)	0.857

Table 3. Quality scores of CAD-related TikTok videos by source. CAD, coronary artery disease; HONcode: Health on the Net Code; PEMAT: Patient Education Materials Assessment Tool; SD, standard deviation.

total DISCERN scores among the four groups ($P = 0.486$) (Table 3). The average score rate of PEMAT was 79.2%, ranging from 75.8 to 79.8% among different creators. No statistically significant differences were observed for PEMAT or HONcode among different creators ($P > 0.05$ for both) (Table 3).

Medical professionals characteristics and video quality

Given that videos posted by medical professionals accounted for the majority of the videos analyzed, we further categorized them into two groups based on their job titles and analyzed the video characteristics and quality between the two groups. Table 4 shows that senior medical professionals posted the majority of the videos ($n = 69$, 70.4%), while those with intermediate titles posted fewer videos ($n = 29$, 29.6%). The videos posted by medical professionals with intermediate titles were longer than those posted by senior medical professionals (131.3 s vs. 107.7 s, $P = 0.024$). The videos posted by senior medical professionals received fewer "likes" (17,120.1 vs. 65,471.0, $P < 0.001$), comments (731.4 vs. 2330.5, $P < 0.001$) and "favorites" (824.2 vs. 6139.6, $P = 0.003$) than those by medical professionals with intermediate titles. In addition, medical professionals with intermediate titles had more followers compared to senior medical professionals (3,140,765.0 vs. 555,500.2, $P < 0.001$) (Table 4).

In Sect. 1 of DISCERN, the videos posted by medical professionals with intermediate professional titles scored higher (29.2 ± 3.3) than those posted by senior medical professionals (28.4 ± 1.8) ($P < 0.001$). HONcode also showed that medical professionals with intermediate professional titles published higher quality videos

	Senior professional title* (*n* = 69)	Intermediate professional title* (*n* = 29)	Total (*n* = 122)	*P* value
Video length (seconds), mean	107.7	131.3	114.4	0.024
Time since posting, mean	12.8	14.2	12.9	0.496
Number of "likes", mean	17120.1	65471.0	31380.2	< 0.001
Number of comments, mean	731.4	2330.5	1202.6	< 0.001
Number of "Favorites", mean	824.2	6139.6	2393.8	0.003
Number of followers of the video creator, mean	555,500.2	3,140,765.0	1,314,824.0	< 0.001

Table 4. CAD-related TikTok videos characteristics by the title of medical professionals. *Senior professional title: a high-level professional designation, equivalent to an attending physician in the U.S. system. Intermediate professional title: a mid-level professional designation, comparable to a fellow physician in the U.S. CAD, coronary artery disease.

	Senior professional title* (*n* = 69)	Intermediate professional title* (*n* = 29)	*P* value
DISCERN			
Reliability of the videos (items 1–8), mean (SD)	28.4 (1.8)	29.2 (3.3)	< 0.001
Quality of treatment choices (items 9–15), mean (SD)	14.1 (6.2)	15.8 (6.5)	0.714
Overall information quality (item 16), mean (SD)	3.2 (0.8)	3.4 (0.9)	0.233
Total DISCERN score, mean (SD)	88.2 (13.2)	93.5 (16.8)	0.120
PEMAT			
Understandability (items 1, 3–5, 8–14, 18–19), mean (SD) [scoring rate (%)]	61.0 (9.7)	58.8 (9.3)	0.816
Actionability (items 20–22, 25), mean (SD) [scoring rate (%)]	20.3 (6.0)	17.4 (6.5)	0.641
Total PEMAT score, mean (SD) [scoring rate (%)]	81.3 (12.1)	76.3 (12.9)	0.705
HonCode			
Total HonCode score, mean (SD)	1.4 (0.5)	1.6 (0.7)	0.006

Table 5. Quality scores of CAD-related TikTok videos by the title of medical professionals. *Senior professional title: A high-level professional designation, equivalent to an attending physician in the U.S. system. Intermediate professional title: A mid-level professional designation, comparable to a fellow physician in the U.S. CAD, coronary artery disease; HONcode: Health on the Net Code; PEMAT: Patient Education Materials Assessment Tool; SD, standard deviation.

(1.6 ± 0.7) than those with senior titles (1.4 ± 0.5) ($P = 0.006$). No statistical significance was found in other scales ($P > 0.05$) (Table 5).

Correlation between PEMAT and DISCERN

Figure 3 analyzed the linear correlation between PEMAT and DISCERN and indicated no statistically significant correlation between the two scales ($P = 0.052$). In simple linear analysis models, a nonlinear correlation was illustrated between DISCERN and PEMAT scores (Fig. 3).

Correlation between video characteristics and quality scores

Figure 4 showed the correlation between video characteristics and the scales. We found that time since posting was negatively correlated with DISCERN ($P = 0.021$) and PEMAT ($P = 0.037$) scores. The number of "favorites" showed a positive linear correlation to DISCERN ($P = 0.007$). Other characteristics showed no significant correlation with either scale ($P > 0.05$) (Fig. 4).

Temporal trends in video quality

Analysis of video quality across different time periods revealed significant differences in DISCERN scores ($P = 0.022$). Videos posted from 2023 to present showed the highest DISCERN scores (93.87 ± 16.13), compared to 2022–2023 (85.87 ± 12.47) and before 2022 (87.59 ± 13.31). For HONcode scores, no significant difference was found across time periods ($P = 0.170$), with all periods showing a median score of 1.00. Similarly, PEMAT scores showed no significant differences across periods ($P = 0.281$), though there was a slight increasing trend from before 2022 (113.38 ± 2.40) to 2023–present (114.63 ± 4.38) (Table 6).

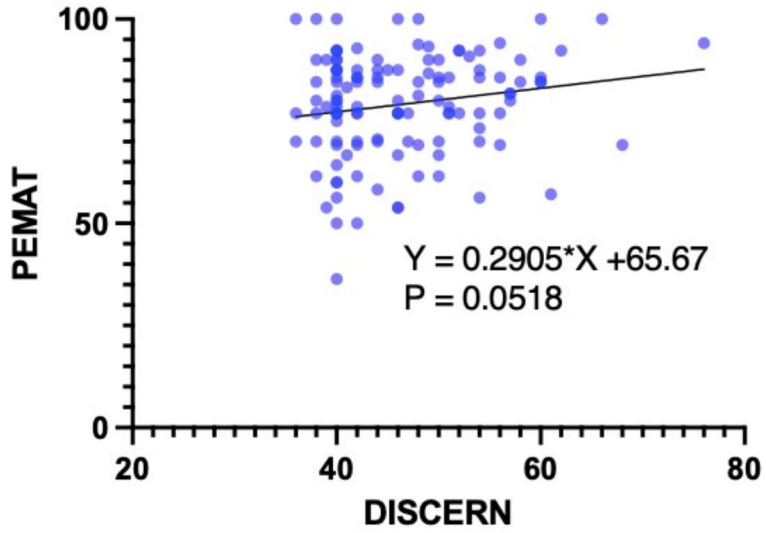

Fig. 3. Correlation analysis between PEMAT and DISCERN scores.

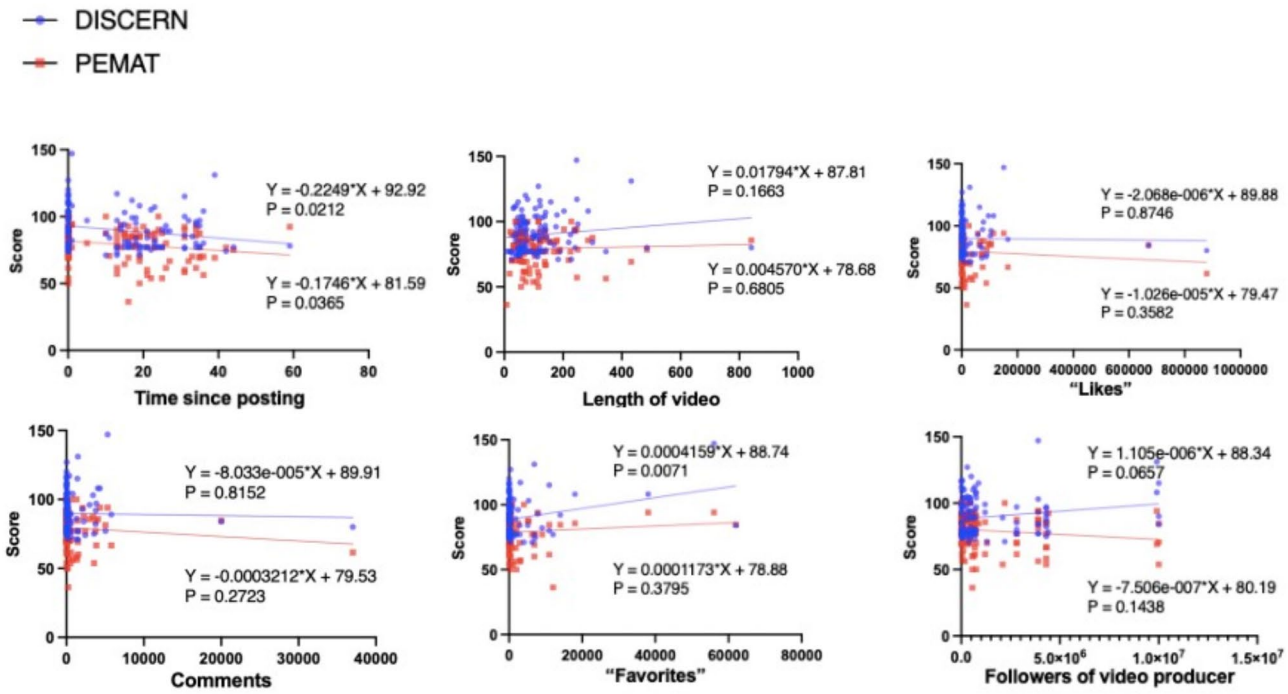

Fig. 4. Correlation analysis between characteristics of video and quality scores.

	2023–present ($n=54$)	2022–2023 ($n=39$)	Before 2022 ($n=29$)	*P* value
Total DISCERN score	93.87 ± 16.13	85.87 ± 12.47	87.59 ± 13.31	0.022
Total HonCode score	1.00 (1.00–2.00)	1.00 (1.00–1.50)	1.00 (1.00–2.00)	0.170
Total PEMAT score	114.63 ± 4.38	113.51 ± 4.47	113.38 ± 2.40	0.281

Table 6. Comparison of the quality scores of CAD-related TikTok videos across time periods. CAD, coronary artery disease; HONcode: Health on the Net Code; PEMAT: Patient Education Materials Assessment Tool.

Discussion

In the past few years, many health-related videos have been published on TikTok, one of the largest short video platforms in China, prompting studies on their characteristics, quality, and sources. Previous studies have suggested that the quality of these health-related TikTok videos is low[15,19]. However, our analysis of 122 CAD-related videos on Chinese TikTok using three scales (DISCERN, PEMAT, and HONcode) yielded different results, with the average quality scores falling in the mid-range and varying across the different scales. Specifically, the average score of the analyzed videos was 46.5/80 for DISCERN and 79.2% for PEMAT. HONcode was found to be less suitable for short video evaluation due to its low score rate (1.4/8). Among the distribution of each sub-item in three scales, DISCERN and PEMAT were relatively uniform. However, the score distribution of HONcode was mainly concentrated in the first two items, while no video scored on the last three items. This discrepancy may be due to the fact that HONcode is most commonly used in evaluating longer videos, focusing on content detail, references, and funding sources, while the DISCERN and PEMAT scales evaluate multiple dimensions, and videos of different lengths were evaluated with corresponding items[13,25–27]. Therefore, HONcode may not be suitable for short video quality evaluation.

Our study also addressed a gap by analyzing score consistency, revealing no correlation between DISCERN and PEMAT. Although both scales may reflect the quality of short videos on TikTok in their respective dimensions of applicability, these dimensions lacked consistency. Therefore, there is an urgent need for a unified and comprehensive short video evaluation scale to assess the quality of short videos. Recent research has provided valuable insights into the quality assessment of health-related short videos. Mao et al. developed the Acute Pancreatitis Content Score (APCS) as a supplementary evaluation tool alongside DISCERN and HONcode, demonstrating a positive correlation between video duration and quality scores (DISCERN: $r = 0.309$, APCS: $r = 0.407$)[20]. A scoping review by Li et al. revealed that most existing assessment tools were not specifically designed for short videos and lacked reliability testing. Many studies also failed to report evaluator identities and reliability across evaluators[28]. These studies point to several directions for future research on assessing the quality of TikTok videos. First, similar to the APCS approach, disease-specific evaluation criteria should be developed to complement general tools like DISCERN to help determine whether a video adequately addresses essential aspects of specific conditions. For CAD-related videos, evaluation should focus on key areas such as risk factors, symptoms, diagnostic methods, treatment options, and prevention strategies. Also, given the importance of evaluator diversity and reliability, as emphasized by Li et al., future research should involve evaluators from diverse backgrounds and systematically report reliability across different evaluators. This approach would enhance the credibility and comparability of quality assessments across studies.

A previous study using DISCERN score showed that health-related videos posted by healthcare professionals are of better quality than those posted by non-healthcare professionals[29]. Building on this, we introduced a novel classification for analyzing videos posted by medical professionals, distinguishing between intermediate and senior professional titles. We found that videos posted by creators with intermediate professional titles were longer, of higher quality, and received more "likes", comments, and "favorites" compared to those posted by senior medical professionals. The higher quality of videos from medical professionals with intermediate titles is likely attributable to their more up-to-date content. Also, medical professionals with intermediate titles had more followers. Several potential factors may contribute to these observations. First, medical professionals with intermediate titles are typically younger than senior professionals, hence may be more familiar with social media platforms and digital communication strategies. Second, medical professionals with intermediate titles may have a better balance between clinical experience and time available for content creation. While possessing sufficient medical expertise, they may have more flexibility in their schedules compared to senior professionals who often have more administrative and clinical responsibilities, allowing them more time to dedicate to social media content creation. Third, their career development stage may motivate them to prioritize establishing a professional reputation through social media presence. However, we acknowledge that these explanations are largely speculative and require further investigation. Future studies could explore this topic through qualitative interviews with medical professionals with different professional titles to understand their social media content creation practices. Additionally, analysis of the time investment in content creation across different professional levels and evaluating the relationship between content quality and creators' social media experience may offer valuable insights into the differences observed in our study.

Previous studies have identified several factors associated with video quality, such as a positive correlation between video length and quality[30], as well as links to creators and the number of "likes"[31,32]. In our analysis of the characteristics of high-quality TikTok videos, we found that video quality was associated with time since posting and the number of "favorites". The time since posting showed a negative correlation with both DISCERN and PEMAT scores in simple linear model, indicating that newer videos tend to have higher quality, likely due to the growing experience and skills of video creators. Temporal analysis also revealed that DISCERN scores improved significantly in 2023–present compared to earlier periods ($P = 0.022$), while HONcode and PEMAT scores remained stable across different time periods. Among engagement metrics, only the number of "favorites" showed a positive correlation with DISCERN scores ($P = 0.007$), while other metrics like the numbers of "likes" and comments showed no significant correlation with quality scores. This finding may be attributed to several factors. First, the "favorites" feature on TikTok serves a distinct purpose from other engagement metrics, as it allows users to save content for future reference, suggesting a more deliberate evaluation of content value by users. For CAD-related videos, this correlation between the number of "favorites" and quality scores may indicate that users, particularly those affected by CAD who are typically older, can identify and actively save high-quality medical information for future reference or sharing with others[33,34]. The lack of correlation between video quality and other engagement metrics ("likes" and comments) suggests that immediate engagement behaviors may be driven by factors beyond content quality, such as video production style, background music, or the creator's presentation skills. This disconnect between quality and popular engagement metrics poses challenges

for both content creators and platform algorithms in promoting high-quality health information. These findings have important implications for health communication on social media platforms. While platforms typically use the number of "likes" and comments as indicators of content value, our results suggest that the number of "favorites" may be a better metric for identifying high-quality health information, particularly for chronic disease-related content. This insight could inform the development of more effective content recommendation algorithms for health-related videos. Based on these findings, we suggest that short video creators should focus on enhancing the depth of the video content and provide more suggestions on prevention and treatment to improve the quality of videos.

In conclusion, our findings have several important practical implications for health communication on social media platforms. For content creators, especially healthcare professionals, our results suggest that investing time in content quality, particularly in aspects measured by DISCERN and PEMAT, can improve user engagement through the "favorites" feature. The higher quality scores of newer videos indicate that creators should stay updated with platform features and audience preferences. For social media platforms, our results suggest that the number of "favorites" may be a more reliable indicator of content quality than other engagement metrics, offering valuable insights for developing algorithms that better promote high-quality health information.

Limitations

Several limitations of this study should be considered. Firstly, the sample size of 122 TikTok videos may be insufficient, and the sample selection was influenced by TikTok's algorithm, potentially introducing bias. Secondly, this study focused on short TikTok videos in China, and the findings may not be applicable to health-related videos on TikTok in other countries due to different healthcare systems, cultural contexts, user behaviors and healthcare content regulations. The content quality standards, user engagement patterns and the ways in which medical professionals create and share information may vary significantly across regions. Lastly, other social media platforms such as YouTube, Instagram, or regional platforms may have different features and algorithms that influence content creation and distribution, which could lead to different quality patterns. Future studies should include a more diverse range of videos from multiple countries and regions, and compare content quality across different platforms to provide more comprehensive insights into health communication on social media.

Data availability

The data supporting the conclusions of this article is available in the TikTok platform (Chinese version, as of July 18, 2023). For data requests or further information, please contact the corresponding author Lin Zhao at trichina2007@126.com.

Received: 14 October 2024; Accepted: 11 March 2025
Published online: 17 March 2025

References

1. Akella, A. & Akella, S. Machine learning algorithms for predicting coronary artery disease: efforts toward an open source solution. *Future Sci. OA*. **7** (6), FSO698 (2021).
2. In China, T. & Hu, S. S. Report on cardiovascular health and diseases in China 2021: an updated summary. *J. Geriatr. Cardiol.* **20** (6), 399–430 (2023).
3. China NHCotPsRo. China health statistics yearbook. : Beijing: Peking Union Medical College Press.; 2021. (2020).
4. Bauersachs, R. et al. Burden of Coronary Artery Disease and Peripheral Artery Disease: A Literature Review. *Cardiovasc Ther*. 8295054 (2019). (2019).
5. Cheng, M. M. et al. Optimizing CAD diagnosis in China with CT angiography. *J. Cardiovasc. Comput. Tomogr.* **3** (3), 153–158 (2009).
6. Cui, Q. et al. Impact of cardiovascular health and genetic risk on coronary artery disease in Chinese adults. *Heart* **109** (10), 756–762 (2023).
7. Zhang, Y. et al. HDL subfractions and very early CAD: novel findings from untreated patients in a Chinese cohort. *Sci. Rep.* **6**, 30741 (2016).
8. Olvera, C., Stebbins, G. T., Goetz, C. G. & Kompoliti, K. TikTok tics: A pandemic within a pandemic. *Mov. Disord Clin. Pract.* **8** (8), 1200–1205 (2021).
9. McCashin, D. & Murphy, C. M. Using TikTok for public and youth mental health - A systematic review and content analysis. *Clin. Child. Psychol. Psychiatry*. **28** (1), 279–306 (2023).
10. Kong, W., Song, S., Zhao, Y. C., Zhu, Q. & Sha, L. TikTok as a health information source: assessment of the quality of information in Diabetes-Related videos. *J. Med. Internet Res.* **23** (9), e30409 (2021).
11. Song, S. 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.* **23** (12), e28318 (2021).
12. Zheluk, A. A., Anderson, J. & Dineen-Griffin, S. Adolescent anxiety and TikTok: an exploratory study. *Cureus* **14** (12), e32530 (2022).
13. Portillo, I. A., Johnson, C. V. & Johnson, S. Y. Quality evaluation of consumer health information websites found on Google using DISCERN, CRAAP, and HONcode. *Med. Ref. Serv. Q.* **40** (4), 396–407 (2021).
14. Vishnevetsky, J., Walters, C. B. & Tan, K. S. Interrater reliability of the patient education materials assessment tool (PEMAT). *Patient Educ. Couns.* **101** (3), 490–496 (2018).
15. Chen, Z., Pan, S. & Zuo, S. TikTok and YouTube as sources of information on anal fissure: A comparative analysis. *Front. Public. Health*. **10**, 1000338 (2022).
16. Yeung, A., Ng, E. & Abi-Jaoude, E. TikTok and Attention-Deficit/Hyperactivity disorder: A Cross-Sectional study of social media content quality. *Can. J. Psychiatry*. **67** (12), 899–906 (2022).
17. Ozsoy-Unubol, T. & Alanbay-Yagci, E. YouTube as a source of information on fibromyalgia. *Int. J. Rheum. Dis.* **24** (2), 197–202 (2021).
18. Furukawa, E. et al. Translation, Cross-Cultural adaptation, and validation of the Japanese version of the patient education materials assessment tool (PEMAT). *Int. J. Environ. Res. Public. Health* **19** (23), (2022).
19. Yang, S., Zhan, J. & Xu, X. Is TikTok a high-quality source of information on thyroid cancer? *Endocrine* **81** (2), 270–276 (2023).

20. Mao, T. et al. Evaluation of TikTok videos on acute pancreatitis: content quality and reliability analysis. *BMC Public. Health*. **24** (1), 1216 (2024).
21. Shepperd, S., Charnock, D. & Why *DISCERN? Health Expect.* **1** (2), 134–135 (1998).
22. Shoemaker, S. J., Wolf, M. S. & Brach, C. The Patient Education Materials Assessment Tool (PEMAT) and User's Guide 2013 [Available from: https://www.ahrq.gov/sites/default/files/publications/files/pemat_guide.pdf
23. Pletneva, N., Cruchet, S., Simonet, M. A., Kajiwara, M. & Boyer, C. Results of the 10 HON survey on health and medical internet use. *Stud. Health Technol. Inf.* (2011). 169 73 – 7.
24. Fullard, A. C., Johnston, S. M. & Hehir, D. J. Quality and reliability evaluation of current internet information regarding mesh use in inguinal hernia surgery using HONcode and the DISCERN instrument. *Hernia* **25** (5), 1325–1330 (2021).
25. Laversin, S., Baujard, V., Gaudinat, A., Simonet, M. A. & Boyer, C. Improving the transparency of health information found on the internet through the Honcode: a comparative study. *Stud. Health Technol. Inf.* **169**, 654–658 (2011).
26. Kaya, E. & Gormez, S. Quality and readability of online information on plantar fasciitis and calcaneal spur. *Rheumatol. Int.* **42** (11), 1965–1972 (2022).
27. Wilkens, F. M. et al. YouTube-videos for patient education in Lymphangioleiomyomatosis?? *Respir Res.* **23** (1), 103 (2022).
28. Li, B. et al. Quality assessment of health science-related short videos on TikTok: A scoping review. *Int. J. Med. Inf.* **186**, 105426 (2024).
29. Evans, E., Gory, L. B. & O'Kane, A. TikTok: an opportunity for antibiotic education?? *Innov. Pharm.* **13** (4), (2022).
30. Sun, F., Zheng, S. & Wu, J. Quality of information in gallstone disease videos on TikTok: Cross-sectional study. *J. Med. Internet Res.* **25**, e39162 (2023).
31. McBriar, J. D. et al. #Neurosurgery: A Cross-Sectional analysis of neurosurgical content on TikTok. *World Neurosurg. X*. **17**, 100137 (2023).
32. Gupta, A. K., Polla Ravi, S. & Wang, T. Alopecia areata and pattern hair loss (androgenetic alopecia) on social media - Current public interest trends and cross-sectional analysis of YouTube and TikTok contents. *J. Cosmet. Dermatol.* **22** (2), 586–592 (2023).
33. Chamberlain, J. D., Bowman, C. R. & Dennis, N. A. Age-related differences in encoding-retrieval similarity and their relationship to false memory. *Neurobiol. Aging*. **113**, 15–27 (2022).
34. Pitts, B. L., Smith, M. E., Newberry, K. M. & Bailey, H. R. Semantic knowledge attenuates age-related differences in event segmentation and episodic memory. *Mem. Cognit*. **50** (3), 586–600 (2022).

Author contributions

Peng Hao: Conceptualization, methodology, data curation, draft writing, project management & manuscript review. Guixin Liu: Resource provision & manuscript review. Shuang Lian: Investigation & manuscript review. Jiaxu Huang: Resource provision & manuscript review. Lin Zhao: Project supervision & manuscript review. All reviewed manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-93986-3.

Correspondence and requests for materials should be addressed to L.Z.

Reprints and permissions information is available at www.nature.com/reprints.

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