



OPEN ACCESS

EDITED BY

Bárbara Mercedes Castillo Abdul Hadi,
University of Design, Innovation and
Technology (UDIT), Spain

REVIEWED BY

Zixuan He,
Second Military Medical University,
China
Laeli Nur Hasanah,
PGRI University of Yogyakarta, Indonesia

*CORRESPONDENCE

Hanyi Yu
✉ yhydtc1@gmail.com

[†]These authors have contributed equally
to this work

RECEIVED 19 November 2025

REVISED 01 March 2026

ACCEPTED 09 March 2026

PUBLISHED 19 March 2026

CITATION

Zhou S, Zhang Y, Chen X, Zhong Q,
Sun F, Xu R, Sun Z, Sun J, Yang L, Xie Z,
Zhang Q, Yan S, Cai Y, Yu H and
Xu Y (2026) Reliability and quality
assessment of internet videos as
guidance for dietary weight loss
intervention: a cross-sectional study in
China.

Front. Public Health 14:1749849.
doi: 10.3389/fpubh.2026.1749849

COPYRIGHT

© 2026 Zhou, Zhang, Chen, Zhong, Sun,
Xu, Sun, Yang, Xie, Zhang, Yan, Cai,
Yu and Xu. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance
with accepted academic practice. No
use, distribution or reproduction is
permitted which does not comply with
these terms.

Reliability and quality assessment of internet videos as guidance for dietary weight loss intervention: a cross-sectional study in China

Shiqi Zhou^{1†}, Yongjian Zhang^{2†}, Xiaohong Chen²,
Qianqian Zhong², Fengqin Sun³, Rui Xu⁴, Zhongli Sun⁴,
Junying Sun⁴, Lin Yang², Zhuoxian Xie¹, Qianle Zhang¹,
Shiqi Yan¹, Yongjiang Cai², Hanyi Yu^{1*} and Yanwu Xu^{1,5}

¹School of Future Technology, South China University of Technology, Guangzhou, China, ²Health Management Center, Peking University Shenzhen Hospital, Shenzhen, China, ³Shenzhen Family Doctors Association, Shenzhen, China, ⁴Shenzhen Nanshan District Medical Group Headquarters, Shenzhen, China, ⁵Pazhou Lab, Guangzhou, China

Background: Obesity, a chronic condition affecting multiple physiological systems, poses major public health challenges. Dietary interventions are widely recognized as effective strategies for weight management. With over 4 billion Internet users worldwide, an increasing number of individuals rely on online platforms for health information. In China, TikTok, Bilibili, and Kwai are major channels for disseminating health-related content. However, the quality of dietary weight loss information on these platforms remains unclear.

Objective: This study aims to assess the reliability and quality of the information in Chinese videos on dietary weight loss shared on the Bilibili, TikTok, and Kwai, three video-sharing platforms.

Methods: We identified the top 100 dietary weight-loss videos on each platform in February 2024, resulting in a total of 300 videos. Video information quality and reliability were assessed using the Global Quality Score (GQS) and modified DISCERN (mDISCERN). Correlations between video quality and video characteristics were also analyzed.

Results: The average GQS scores for Bilibili, TikTok, and Kwai were 2.04, 1.81, and 1.70, respectively, while the average mDISCERN scores were 2.01, 1.81, and 1.73. Median scores for both tools across all platforms were 2. Bilibili showed significantly higher GQS scores than TikTok and Kwai ($p < 0.01$ and $p < 0.05$, respectively). Regarding mDISCERN, Bilibili scored significantly higher than TikTok ($p < 0.05$), while the difference with Kwai was not statistically significant ($p = 0.08$). Nevertheless, none of the platforms achieved scores above 3, indicating generally low information quality and reliability. Significant positive correlations were found between video duration and both GQS ($r = 0.41$, $p < 0.01$) and mDISCERN ($r = 0.32$, $p < 0.01$). Additionally, strong correlations were observed between likes and saves ($r = 0.90$, $p < 0.01$), likes and comments ($r = 0.92$, $p < 0.01$), and saves and comments ($r = 0.86$, $p < 0.01$).

Conclusion: While acknowledging limitations regarding cross-sectional design and specific sampling of single keyword and Chinese platforms, our findings highlight that the prevalence of low-quality videos on Chinese social media exposes viewers to significant risks of misinformation and inappropriate dieting. These findings underscore the need to promote digital health literacy and improve strategies for digital health communication.

KEYWORDS

dietary intervention, internet videos, quality assessment, reliability assessment, weight loss

1 Introduction

Obesity is a chronic disease, like hypertension and asthma, adversely effects affecting nearly on all physiological functions of the human body. The World Health Organization defines overweight and obesity as abnormal or excessive fat accumulation that poses a threat to health (1). About 2000 years ago, people began to pay attention to the impact of obesity on incidence and mortality rates (2). A substantial body of research has demonstrated that obesity increases the risk of developing diabetes (3), cardiovascular diseases (3, 4), or psychological disorders such as depression (5). Obesity is a serious public health threat. The core principle of any obesity treatment is to alter the balance between energy intake and energy expenditure, ensuring that energy expenditure exceeds energy intake. While medicinal interventions and bariatric surgery are common weight loss methods (6), they often require substantial resources and can be costly or inefficient for general population. However, a growing body of evidence suggests that dietary intervention is a critical factor in short-term and long-term weight loss (7–9), and can achieve ideal weight management without compromising health. Therefore, it is crucial for the public to assess the reliability of dietary intervention information gathered in daily life for weight loss.

Technological advancements have revolutionized the way people access medical information has undergone significant changes. Compared to traditional methods such as consulting doctors or reading medical books, more and more individuals prefer to seek medical information online (10, 11). As of 2021, the number of global Internet users has exceeded 4 billion (12), and studies have found that one-third of American adults use the Internet to learn about health issues (13). In recent years, video information has become increasingly popular. Statistics show that videos with the hashtag #weightloss on TikTok have been viewed more than 9.7 billion times worldwide (14), indicating that healthy weight loss is a global hot topic. Compared to traditional text, social media videos offer key advantages (15, 16). First, visual information is more easily absorbed and remembered than text, allowing for the rapid and cost-effective dissemination of health information to a broad audience. Second, the rich visual effects of videos can encourage users to spontaneously adopt healthy behaviors. However, due to the large number of content creators on video social platforms and less regulation of video content, the quality of online information on diet and weight loss varies significantly. Driven by the desire for engagement metrics such as likes and saves, creators often exaggerate effects to attract viewers. Consequently, this prioritization of popularity over accuracy means audience may inadvertently encounter incomplete or misleading information (17, 18). Therefore, quality assessment of medical science popularization videos online is essential to mitigate online misinformation and promote public health literacy.

In China, Bilibili, TikTok (Chinese version also named Douyin), and Kwai (Chinese version also named Kuai Shou) are three most popular video-sharing platforms. These platforms are currently the main online media for disseminating health information online, with a monthly activity level exceeding 100 million people (19). All

information on these platforms is freely accessible and users only need to search for keywords of interest to view related videos. Previous studies have evaluated the quality of popular medical science videos on different topics on TikTok and Bilibili. Videos related to plastic surgery are considered reliable (20), while the quality of videos on liver cancer and gallstones is generally unsatisfactory (21, 22). Their findings indicate that the overall quality of medical popular science videos on these topics remains suboptimal. We also found that several studies have specifically analyzed the quality of weight loss videos on English-language platforms such as YouTube and TikTok (23, 24). However, to the best of our knowledge, there has been no research analyzing the quality and reliability of videos dietary weight loss intervention on Chinese video-sharing platforms. This gap is particularly concerning given the unique short-video ecology in China, where rapid algorithmic dissemination often outpaces content verification. In clinical practice, this specific digital environment has been increasingly linked to patients adopting unsafe weight-loss behaviors based on trending videos. In response to these challenges, this study aims to assess the quality of the most popular dietary weight loss videos on Bilibili, TikTok, and Kwai, providing an empirical basis for advancing digital health communications and enhancing public health literacy.

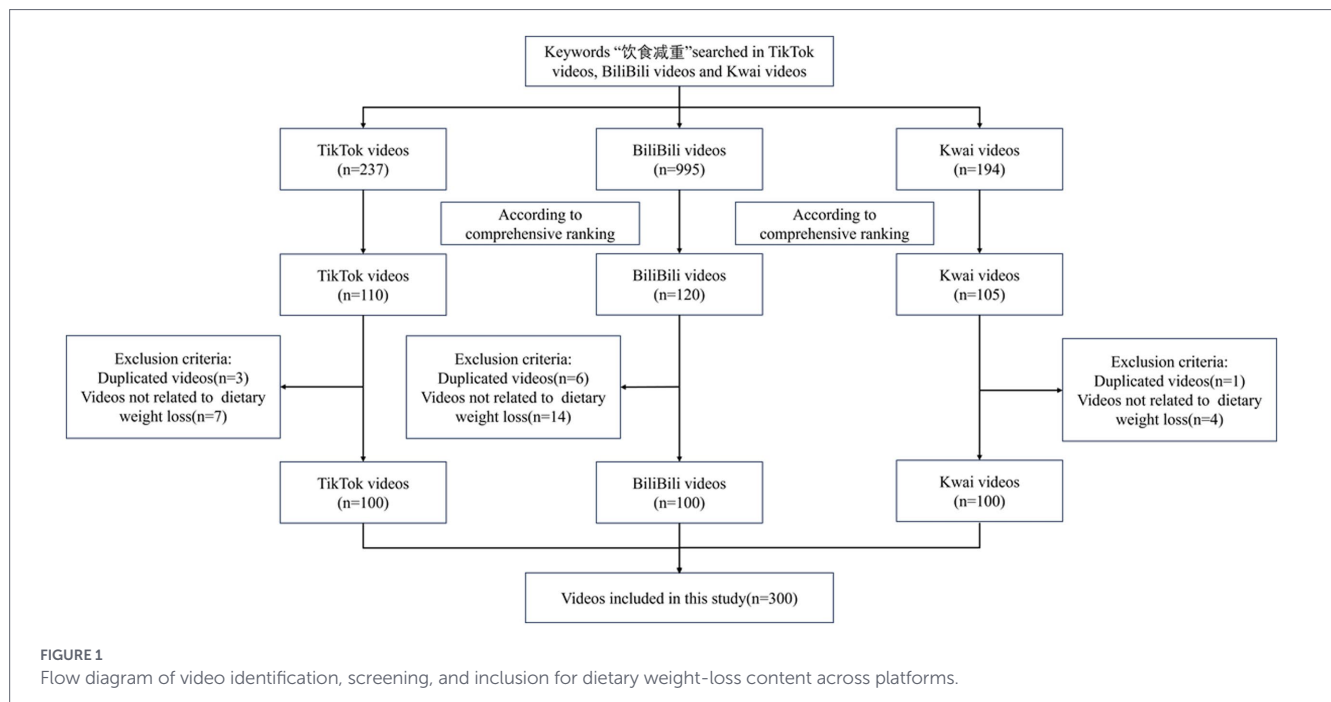
2 Method

2.1 Search strategy and data collection

As shown in Figure 1, in this cross-sectional study, we searched for videos related to “dietary weight loss” (饮食减重) on Bilibili (Chinese version 7.68.0), TikTok (Chinese version 28.6.0), and Kwai (Chinese version 12.0.40) on February 28th, 2024. For conceptual clarity, “dietary weight loss” was operationally defined as weight reduction strategies primarily achieved through modifications of dietary intake, including caloric restriction, macronutrient composition changes (e.g., ketogenic, low-carbohydrate, or low-fat diets), meal replacements, and time-restricted eating (e.g., intermittent fasting). Videos focusing primarily on exercise-only interventions, bariatric surgery, or prescription weight-loss medications were excluded.

To mitigate biases in video recommendations stemming from the personalized algorithms of the platforms, we conducted searches using newly created accounts on each platform without imposing any additional constraints. Videos were initially retrieved and ranked according to the platforms’ default recommendation algorithms. Subsequently, we screened the search results to exclude duplicate content and retained all videos related to the topic of dietary weight loss. Videos identified as part of a continuous serial topic were aggregated as a single one to ensure the completeness of the information provided. In accordance with established practices in recent digital health research, the top 100 eligible videos from each platform were selected to form the final dataset (21, 22).

To ensure the credibility of the video content, we verified the details of the video publishers. For instance, if a publisher was purported to be a medical professional from a hospital, we confirmed the



accuracy of this information on the hospital's official website. Beyond this, we also gathered video metadata for the videos, including the title, the publisher's name, duration, and engagement metrics (the count of likes, saves, and comments). All collected data were securely stored in Excel (Microsoft Inc.) for further analysis.

To facilitate comparative analysis of information quality, we classified the videos according to two key dimensions. First, we categorized the videos into two groups based on their publishers: professional doctors and non-doctors. Second, the videos were classified into four content types: personal experiences, weight loss knowledge, advertisements, and other categories.

2.2 Ethical considerations

This study did not utilize any clinical data, human samples, or laboratory animals. All data were sourced from publicly available videos on Bilibili, TikTok, and Kwai, ensuring no personal privacy issues were involved. As the study did not engage directly with users, it did not require ethical review.

2.3 Video assessment

2.3.1 Assessment criteria

We utilized the mDISCERN and the Global Quality Score (GQS) tools to assess the reliability and informational quality of the videos, respectively, the validity of which has been confirmed in previous studies (25–27). The mDISCERN is an adaptation of the DISCERN instrument proposed by Singh et al., and it evaluates the reliability of videos across five dimensions: clarity, reliability, fairness, references, and rigor (28). Each dimension is scored individually, with a total score ranging from 0 to 5; the higher the score, the better the reliability of the video. Compared to DISCERN, mDISCERN demonstrates improved convenience and accuracy in assessing video quality. The GQS, introduced by Bernard et al., was initially designed to evaluate the quality of health information on websites and has since been

widely adopted for assessing the quality of video information (29). The GQS systematically evaluates the quality of videos based on informational quality, traffic, and usefulness, with a scoring range from 1 (very poor) to 5 (excellent).

2.3.2 Video assessment procedure

We compiled the collected original videos and pertinent information, and then dispatched them to the professional raters involved in the scoring process. All participating video raters had extensive clinical expertise in dietary weight reduction. To minimize scoring discrepancies, the mDISCERN and GQS instruments were translated into Chinese to better align with the linguistic and contextual characteristics of Chinese short-video platforms. The detailed scoring criteria of both instruments, along with their corresponding Chinese translations, are presented in [Supplementary Tables S1, S2](#), respectively. Subsequently, every rater thoroughly reviewed the Chinese-version criteria for GQS and mDISCERN. Prior to formal evaluation, a pilot scoring session was conducted on a subset of 10 videos to ensure a consistent understanding of the rating guidelines. For vlogs or humorous content containing personal narratives, raters were instructed to focus strictly on the extractable health assertions, disregarding entertainment elements. For content involving mixed languages or regional dialects, standard Chinese subtitles served as the primary reference for accuracy to ensure consistency among raters.

To prevent bias, a strict blinding protocol was enforced: platform watermarks were blurred, and raters were blinded to all engagement metrics and origin sources. Each video was then initially assessed by three primary raters (JS, QZ, and ZS). If their scores were unanimous, that score was final. For scores within a one-point range, the principle of majority determined the final score (e.g., mDISCERN scores of 1, 2, 2 and GQS scores of 3, 3, 2 would result in mDISCERN = 2, DQS = 3). If the score discrepancy exceeded one point (e.g., 2, 3, 4), two secondary raters (FS and RX) reassessed the video. If discrepancies remained, a tertiary grader (YZ) then held a final discussion to reach a consensus. Rat

recorded their scores in a scoring spreadsheet; primary raters entered scores for all videos, while secondary and tertiary raters only recorded scores for videos they reviewed.

To further explore the difference in video ratings between professional physicians and the general audience, we engaged three raters (QL, SY, and ZX) with higher education but not medical expertise. They underwent the same standardized training and pilot scoring process to rate the videos based on the same criteria.

2.4 Statistical analysis

In this study, Data normality was first assessed using the Shapiro–Wilk test. As the data were not normally distributed, continuous variables are presented as medians with interquartile ranges (IQR), and categorical variables are presented as frequencies and percentages. Differences in continuous variables between two independent groups (e.g., doctor vs. non-doctor) were assessed using the Mann–Whitney U test.

To evaluate the reliability of the video assessments, we calculated the Fleiss' Kappa coefficient to measure inter-rater agreement among the three primary raters. Additionally, to assess the divergence between professional and non-professional perspectives, Cohen's Kappa coefficient was computed between the final scores of the two rater groups. A kappa value greater than 0.60 was considered indicative of substantial agreement.

Based on prior literature, we hypothesized *a priori* that longer video duration would be positively associated with quality scores, while engagement metrics (likes, saves, and comments) might show negative or non-significant associations. Accordingly, Spearman's rank correlation analysis was utilized to examine the bivariate relationships between video quality scores (GQS, mDISCERN) and video characteristics.

To further identify independent determinants of video quality and reliability while controlling for confounding factors, multivariable linear regression analyses were performed separately for each platform. Prior to the regression analysis, categorical variables were recoded into binary dummy variables to ensure statistical validity. Uploader type was coded as a binary variable (1 = professional doctor, 0 = non-doctor). Similarly, given the limited sample size in certain subgroups, content category was reclassified into a binary variable: 'Weight Loss Knowledge' (coded as 1) versus 'Other Content' (including personal experiences, advertisements, and others, coded as 0).

All statistical analyses were performed using Python with the `scipy` and `statsmodels` libraries. Two-sided *p*-values < 0.05 were considered statistically significant. Given the exploratory nature of this study, we did not apply adjustments for multiple comparisons.

Therefore, the multiple statistical tests performed may increase the risk of Type I errors (false positives) and marginal *p*-values should be interpreted with caution. For the regression results, coefficients (β), standard errors (SE), 95% confidence intervals (CIs), adjusted R^2 , and variance inflation factors (VIFs) were reported to evaluate model fit.

3 Results

3.1 Basic characteristics of videos

We used the keyword “dietary weight loss” to retrieve videos on three platforms, respectively. After the screening, the top 100 videos in Bilibili, TikTok, and Kwai were retained. Table 1 summarizes the descriptive statistics of the video characteristics, and Figure 2 visually compares the data distributions across the three platforms using box plots. The video duration on Bilibili was significantly longer than that on TikTok and Kwai ($p < 0.01$). Specifically, the median duration for Bilibili was 325 s (IQR: 385 s), whereas TikTok and Kwai featured much shorter formats, with medians of 72 s (IQR: 65 s) and 70 s (IQR: 64 s), respectively. In contrast, TikTok and Kwai exhibited significantly higher engagement metrics (likes, saves, and comments) compared to Bilibili ($p < 0.01$), likely due to their shorter video formats. For likes, Kwai (Median: 108.51 k, IQR: 175.25 k) and TikTok (Median: 60.52 k, IQR: 143.25 k) far exceeded Bilibili (Median: 1.85 k, IQR: 3.35 k). A similar trend was observed for saves, where Kwai (Median: 34.03 k, IQR: 58.75 k) and TikTok (Median: 9.49 k, IQR: 26.03 k) were higher than Bilibili (Median: 1.15 k, IQR: 2.86 k). Regarding comments, Kwai videos had a median of 4.62 k (IQR: 6.85 k) and TikTok videos had a median of 2.20 k (IQR: 5.56 k), whereas Bilibili videos had a median of 0.23 k (IQR: 0.34 k).

Figure 3 illustrates the distribution of videos by the uploader and content on the three platforms. Regarding sources, Kwai and Bilibili featured a higher number of videos from professional doctors (20 and 18, respectively) compared to TikTok (5). In terms of content, videos were classified into four categories: personal experiences, weight loss knowledge, advertisements, and others. “Weight loss knowledge” was the dominant category across all platforms (Bilibili: 49; TikTok: 48; Kwai: 56). Notably, TikTok had the highest volume of “personal experience” videos, reaching 50, which significantly exceeded the 30 on Bilibili and 27 on Kwai. Advertisements were also more frequent on Kwai and Bilibili, with 15 and 9 videos respectively, whereas TikTok contained only 2.

TABLE 1 Characteristics of dietary weight loss videos collected from the three platforms.

| Variable | Bilibili | TikTok | Kwai |
|--------------|--------------------|----------------------|-----------------------|
| GQS | 2 (0–4) | 2 (1–3) | 2 (1–3) |
| mDISCERN | 2 (0–5) | 2 (0–3) | 2 (0–3) |
| Duration (s) | 325 (12–13,200) | 72 (6–384) | 70 (6–474) |
| Like (k) | 1.85 (0.04–378.66) | 60.52 (0.04–1765.23) | 108.51 (2.42–1252.89) |
| Save (k) | 1.15 (0.00–5.94) | 9.50 (0.00–70.18) | 34.21 (1.33–441.14) |
| Comment (k) | 0.23 (0.01–4.42) | 2.24 (0.02–207.85) | 4.62 (0.02–84.23) |

Data are presented as median (range). Group differences were evaluated using the Mann–Whitney U test, with $p < 0.05$ considered statistically significant.

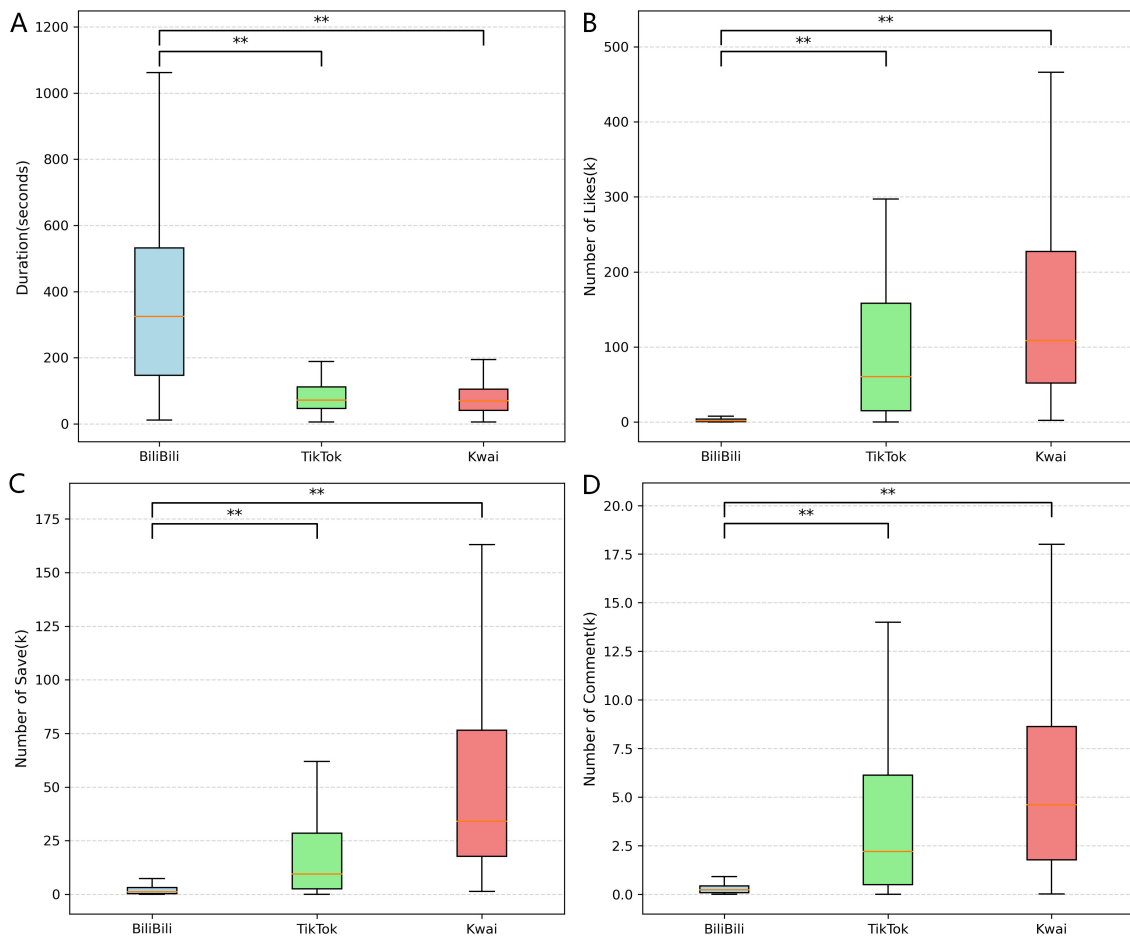


FIGURE 2

Boxplots of dietary weight loss videos characteristics across platforms. Panels show (A) Duration, (B) Likes, (C) Saves, and (D) Comments. In each boxplot, the center line indicates the median, boxes indicate the interquartile range (IQR, 25th–75th percentiles), and whiskers indicate $1.5 \times \text{IQR} / (\text{min} - \text{max})$. Pairwise comparisons between platforms were performed using Mann–Whitney U tests and p -values were unadjusted for multiple comparisons. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.2 Video quality and reliability assessments

3.2.1 Overall comparison across platforms

As shown in Figure 4, after three rounds of rating, the average GQS scores for dietary weight loss videos on Bilibili, TikTok, and Kwai were 2.04, 1.81, and 1.70, respectively, while the average mDISCERN scores were 2.01, 1.81, and 1.73. The median scores for both GQS and mDISCERN across the three platforms were 2. Although Bilibili scored higher than those of Kwai and TikTok (GQS: $p < 0.01$, $p = 0.02$; mDISCERN: $p = 0.02$, $p = 0.08$), overall scores remained below 3 on the three platforms. According to the definition of GQS in Supplementary Table S2, a score of 5 indicates excellent quality, 4 indicates good quality, 3 indicates moderate quality, and 1–2 indicate poor quality. Therefore, the observed mean and median GQS values suggest that low-quality videos were predominant across all platforms. This result may be attributed to insufficient oversight of the scientific rationality and authenticity of their video content on these platforms.

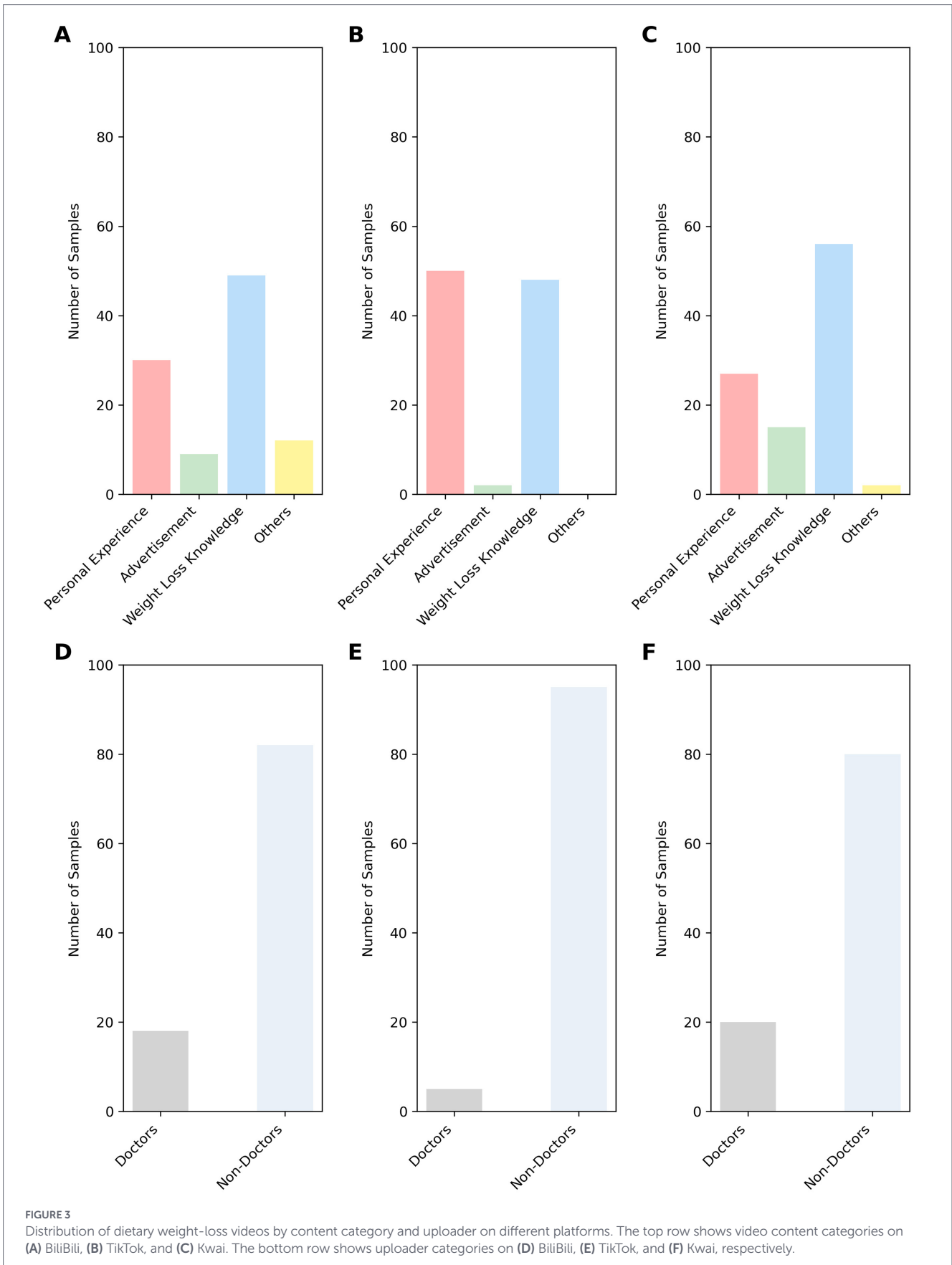
To validate the consistency of our assessment, we analyzed the scoring agreement among the three primary professional raters during the first round of evaluation. Table 2 presents the mean and median scores assigned by each rater, alongside the corresponding agreement metrics. Notably, the analysis revealed exceptional consistency in the

judgments of the professional raters, with all Fleiss Kappa coefficients exceeding 0.90 for both GQS and mDISCERN scores. This near-perfect agreement confirms the robustness and reliability of the quality assessment data used in this study.

3.2.2 Comparison between professional and non-professional raters

Table 3 compares the average GQS and mDISCERN scores assigned by professional and non-professional raters across the three platforms. The results indicated poor agreement between experts and the general audience regarding the quality and reliability of dietary weight loss videos, with the highest Cohen's Kappa coefficient being less than 0.2.

To further characterize the sources of this divergence, we examined agreement across individual dimensions of the mDISCERN scale. For Bilibili videos, the highest agreement was observed in the Clarity dimension (Cohen's $\kappa = 0.43$), whereas the greatest disagreement occurred in Fairness (Cohen's $\kappa = -0.02$). On TikTok, the highest agreement was also found in the Clarity dimension (Cohen's $\kappa = 0.14$), while the largest divergence was observed in the Reliability dimension (Cohen's $\kappa = -0.09$). For Kwai, agreement remained consistently low across all dimensions, with the highest Cohen's κ reaching only 0.05



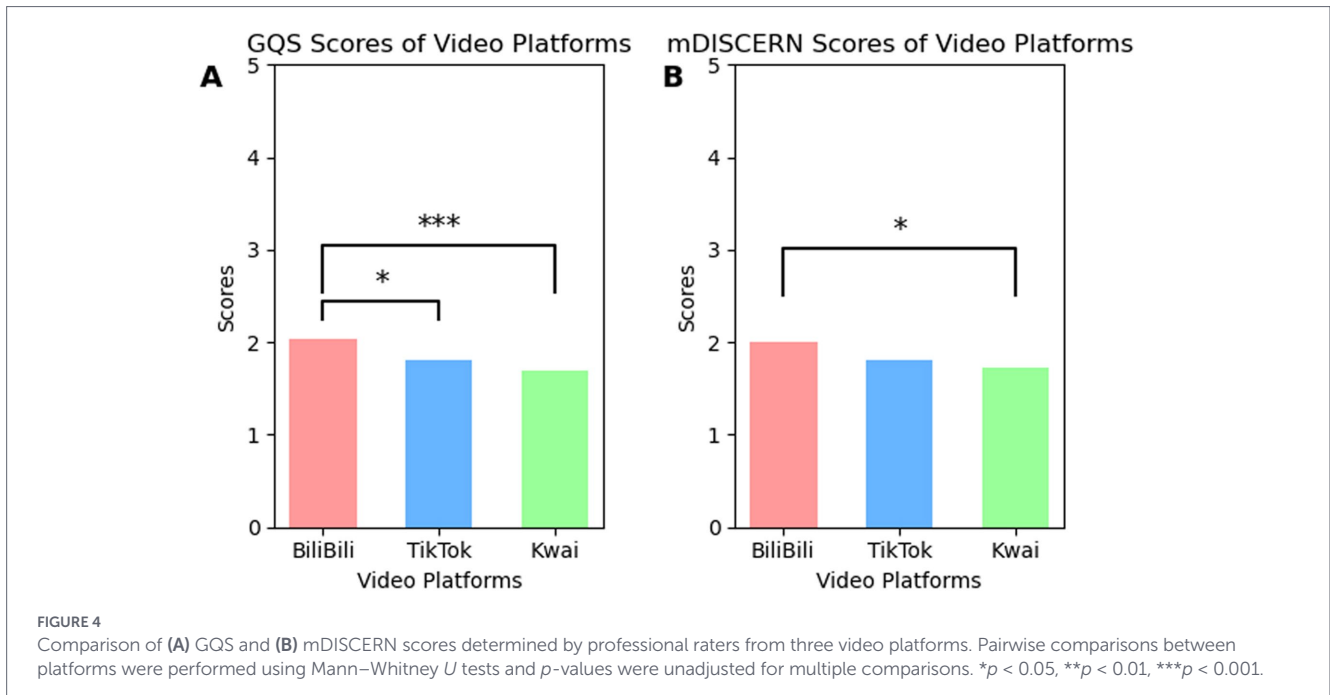


TABLE 2 Comparisons of first-round scores and inter-rater agreement among three professional raters, with Fleiss' Kappa analysis.

| Rater | GQS | | | mDISCERN | | |
|--------------|----------|----------|----------|----------|----------|----------|
| | Bilibili | TikTok | Kwai | Bilibili | TikTok | Kwai |
| Rater#1 | 2.08 (2) | 1.90 (2) | 1.87 (2) | 1.95 (2) | 1.81 (2) | 1.86 (2) |
| Rater#2 | 2.26 (2) | 1.84 (2) | 1.63 (2) | 2.18 (2) | 1.84 (2) | 1.73 (2) |
| Rater#3 | 2.32 (2) | 2.02 (2) | 1.63 (2) | 2.24 (2) | 2.48 (3) | 2.17 (2) |
| Fleiss Kappa | 0.90 | 0.95 | 0.98 | 0.90 | 0.93 | 0.94 |

Data are presented as Mean (Median).

TABLE 3 Comparison of scores between professional and non-professional raters, with Cohen's Kappa analysis.

| Rater type | GQS | | | mDISCERN | | |
|------------------|----------|----------|----------|----------|----------|----------|
| | Bilibili | TikTok | Kwai | Bilibili | TikTok | Kwai |
| Professional | 2.04 (2) | 1.81 (2) | 1.70 (2) | 2.01 (2) | 1.81 (2) | 1.73 (2) |
| Non-professional | 2.10 (2) | 2.29 (2) | 2.77 (2) | 1.67 (2) | 1.91 (2) | 2.41 (3) |
| Cohen's Kappa | 0.19 | 0.05 | 0.04 | 0.13 | 0.12 | 0.03 |

Data are presented as Mean (Median).

(Reliability) and the lowest being -0.02 (References and Rigor). Detailed results are presented in [Supplementary Table S3](#).

This suggests that raters tend to reach higher agreement when assessing whether a video clearly conveys its intended message (Clarity), whereas evaluating whether the content is balanced, accurate, and evidence-based (Reliability or Fairness) requires more professional expertise, leading to lower agreement. We also reviewed a subset of highly discordant videos on Bilibili for qualitative analysis, whose results supported this observation. Specifically, these videos frequently featured popular dietary regimens, such as the 'ketogenic diet,' emphasizing rapid weight loss effects while omitting potential adverse effects or medical contraindications. Non-professional raters, lacking the medical background to detect these omissions, tended to rate such content highly based on its persuasive narrative and promised efficacy.

3.2.3 Comparison across different categories

As shown in [Table 4](#), the relationship between uploader type and video quality differed markedly across platforms. On Bilibili, videos published by professional doctors were rated significantly lower than those by non-doctors in both GQS (1.56 vs. 2.13, $p < 0.05$) and mDISCERN (1.50 vs. 2.10, $p < 0.05$). In contrast, on TikTok, doctor-uploaded videos consistently received higher quality and reliability scores than non-doctor videos, particularly in mDISCERN (2.60 vs. 1.77; $p < 0.05$). On Kwai, the pattern was less consistent: doctor videos scored slightly lower than non-doctor videos on GQS (1.65 vs. 1.71, $p = 0.68$) but demonstrated higher reliability as reflected by mDISCERN scores (2.05 vs. 1.65; $p < 0.05$). These contrasting patterns indicate substantial platform-specific

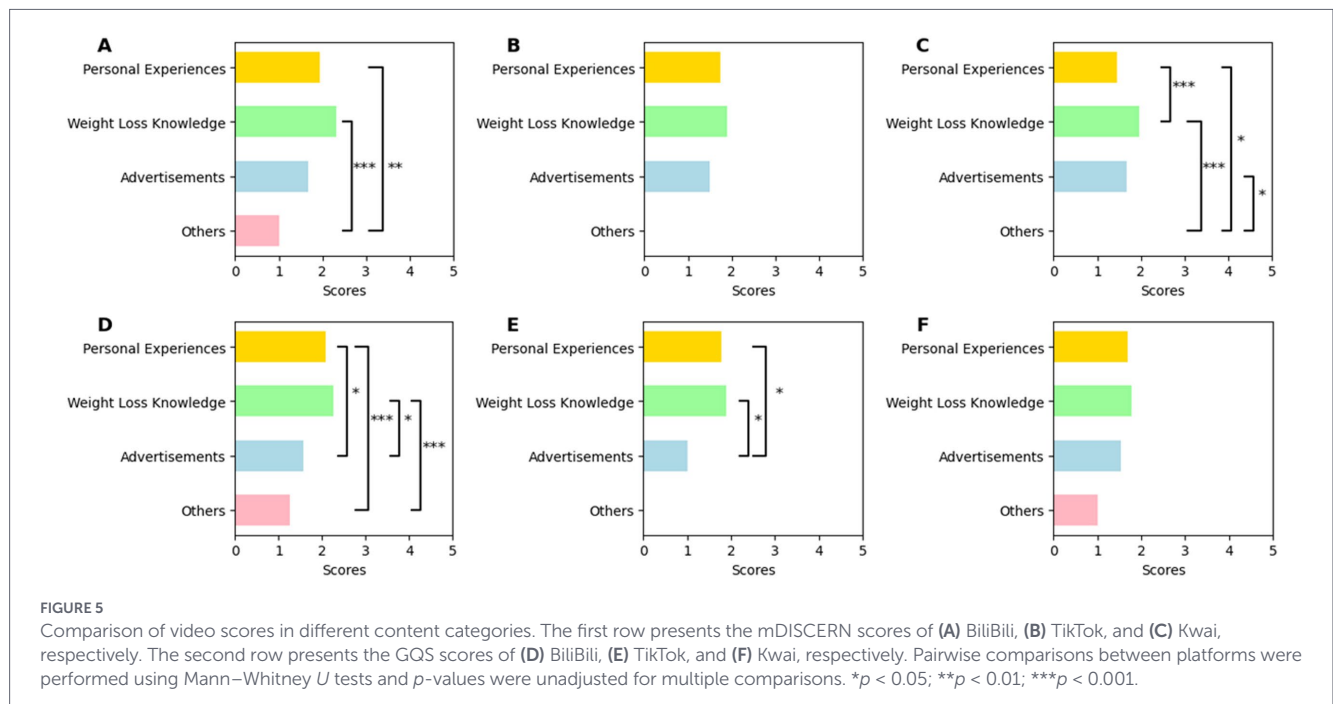
TABLE 4 Comparison of scores for videos from doctor and non-doctor creators across the three platforms.

| Uploader type | GQS | | | mDISCERN | | |
|---------------|----------|--------|------|----------|--------|------|
| | BiliBili | TikTok | Kwai | BiliBili | TikTok | Kwai |
| Doctor | 1.56 | 2.00 | 1.65 | 1.50 | 2.60 | 2.05 |
| Non-doctor | 2.13 | 1.80 | 1.71 | 2.10 | 1.77 | 1.65 |

TABLE 5 Comparison of score for videos with different contents.

| Content type | GQS | | | mDISCERN | | |
|-----------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | BiliBili | TikTok | Kwai | BiliBili | TikTok | Kwai |
| Personal experiences | 2.07 | 1.78 | 1.70 | 1.93 | 1.74 | 1.44 |
| Weight loss knowledge | 2.26 | 1.88 | 1.77 | 2.31 | 1.90 | 1.95 |
| Advertisements | 1.56 | 1.00 | 1.53 | 1.67 | 1.50 | 1.67 |
| Others | 1.25 | / | 1.00 | 1.00 | / | 0.00 |

The highest score of each column is in bold. “/” means that no samples in the corresponding category.



differences in the quality performance of dietary weight loss videos uploaded by doctors.

Table 5 shows that videos focused on weight loss knowledge received the highest GQS and mDISCERN scores across all platforms. Personal experience videos ranked second, except for the mDISCERN scores on Kwai, where advertisement videos held the second-highest position. As TikTok did not have videos in the other category, the corresponding items in the table are marked by “/.” Figure 5 visualizes the comparison among four content categories with significance. Although weight loss knowledge videos were highest in all groups, it had no significant difference from the second place.

3.3 Correlation analysis

Consistent with our *priori* hypotheses, video duration was positively correlated with quality scores, whereas engagement metrics

generally showed negative associations. Positive correlations were observed between GQS scores and video duration ($r = 0.41, p < 0.01$), as well as between mDISCERN scores and video duration ($r = 0.32, p < 0.01$). Strong positive correlations were also found between the number of likes and saves ($r = 0.90, p < 0.01$), likes and comments ($r = 0.92, p < 0.01$), and saves and comments ($r = 0.86, p < 0.01$). In contrast, negative correlations were observed between GQS scores and the number of likes, saves, and comments ($r = -0.18, p < 0.01$; $r = -0.10, p = 0.09$; and $r = -0.17, p < 0.01$, respectively). Similarly, mDISCERN scores were negatively correlated with likes, saves, and comments ($r = -0.12, p < 0.05$; $r = -0.02, p = 0.77$; and $r = -0.13, p < 0.05$). Video duration was also negatively correlated with the number of likes, saves, and comments ($r = -0.36, r = -0.22$, and $r = -0.32$, respectively; all $p < 0.01$).

In addition, we also conducted multivariate regression analyses including video duration, engagement metrics, uploader type, and

content category as predictors. Overall, the associations between video characteristics and quality scores varied across platforms. [Supplementary Table S4](#) presents regression results for Bilibili. Both GQS and mDISCERN scores were negatively associated with uploader type (GQS: $\beta = -0.72$, $p < 0.001$; mDISCERN: $\beta = -0.76$, $p < 0.01$) but positively associated with content category (GQS: $\beta = 0.55$, $p < 0.001$; mDISCERN: $\beta = 0.67$, $p < 0.001$), with longer videos also showing higher mDISCERN scores ($\beta = 0.21$, $p < 0.05$). The adjusted R^2 for GQS and mDISCERN were 0.21 and 0.24, respectively. In contrast, as shown in [Supplementary Table S5](#), video duration on TikTok was positively associated with GQS scores ($\beta = 0.23$, $p < 0.001$), while mDISCERN scores were positively associated with both uploader type ($\beta = 0.75$, $p < 0.01$) and video duration ($\beta = 0.11$, $p < 0.05$). The adjusted R^2 for GQS and mDISCERN were 0.15 and 0.11, respectively. For Kwai, [Supplementary Table S6](#) indicates that GQS scores increased with video duration ($\beta = 0.20$, $p < 0.01$) and the number of saves ($\beta = 0.19$, $p < 0.05$), but decreased with the number of likes ($\beta = -0.30$, $p < 0.01$). Meanwhile, mDISCERN scores on Kwai were positively associated with uploader type ($\beta = 0.34$, $p < 0.05$), content category ($\beta = 0.43$, $p < 0.01$), and video duration ($\beta = 0.14$, $p < 0.05$). The adjusted R^2 for GQS and mDISCERN were 0.14 and 0.24, respectively. Although strong correlations among likes, saves, and comments, only likes (VIF = 7.18) and comments (VIF = 6.86) on TikTok showed moderate multicollinearity.

4 Discussion

4.1 Principal findings

In this cross-sectional study, we collected and reviewed videos on the topic of dietary weight loss from the three most popular video platforms in China, namely Bilibili, TikTok, and Kwai. Despite Bilibili achieving statistically higher scores than TikTok and Kwai, the mean scores across all platforms remained below 3.0, indicating a predominance of low-quality content. Our multivariate analysis revealed three critical trends. First, a trade-off exists between quality and engagement: video duration was a positive predictor of quality but was negatively correlated with popularity metrics, suggesting that algorithms algorithm-driven popularity often favors shorter, less reliable content. Second, the impact of professional sources was platform-dependent; surprisingly, doctor uploader status was positively associated with quality on TikTok but negatively associated on Bilibili. Third, there was poor agreement between professional and lay raters (Cohen's $\kappa < 0.2$). While lay audiences could assess content clarity, they struggled to distinguish narrative appeal from scientific reliability, highlighting a significant gap in digital health literacy.

4.2 Video source, video content and video quality

First, we observed a substantial divergence between professional and non-professional raters across the three platforms. For professional raters, the GQS and mDISCERN scores were ranked from highest to lowest as Bilibili, TikTok, and Kwai. In contrast, non-professional raters ranked the platforms from highest to lowest as Kwai, TikTok, and Bilibili. Our dimension-specific analysis ([Supplementary Table S3](#)) implies that the discrepancy likely stems from Fairness and Reliability.

This low agreement underscores a critical health communication issue: a fundamental disconnect exists between “expert-driven assessment” and “user-perceived value”. Therefore, video creators should be aware of the urgent need for health communication strategies that go beyond content accuracy to address how users actually consume and value information.

Second, we found that videos from Bilibili published by professional doctors scored lower on both the GQS and mDISCERN scales compared to those by non-doctors, which is contrary to previous research findings (19, 21). This divergence may be attributed to topic specificity and information delivery styles. Unlike disease-related topics requiring a strong professional background, dietary weight loss is a lifestyle-integrated topic where non-professional creators often employ narrative-based communication. By blending personal weight-loss journeys with practical tips, these creators enhance the “utility” and “flow” of information, making the videos more accessible to audience. In contrast, videos by professional doctors are more didactic, lacking the same personalized appeal. Moreover, Bilibili videos were substantially longer than those on TikTok and Kwai, which likely increases the demand for the narrative organization ability. If doctors are less experienced in digital content delivery, these limitations could be more salient in long-duration videos.

Third, regarding content categories, videos focusing on “Weight Loss Knowledge” outperformed other three categories across all platforms. Personal experiences with dietary weight loss might spread misinformation, such as losing weight by eating only a cucumber at night or not eating at all and how much weight was lost as a result. In contrast, videos sharing weight loss knowledge were based on scientific evidence, such as the “16 + 8 diet” or “ketogenic diet,” naturally scoring higher than other types of content.

4.3 Comparison with prior research frameworks

In the evolving landscape of social media health analysis, GQS and mDISCERN have emerged as mainstream instruments. Although originally developed for text-based websites, GQS was selected for its holistic focus on “flow,” “ease of understanding,” and “overall usefulness”—qualities well aligned with the rapid, intuitive consumption of short videos. Specifically, GQS has been widely adopted in recent assessments of short videos regarding immune-mediated inflammatory disease (19, 26, 30), gallstone disease (22), cancer (21, 27, 31), and bariatric surgery (23, 32). DISCERN is also a frequently-used instrument in medical video assessment task (20, 22, 24, 26, 33). However, its 16-item structure is often considered too lengthy for rapid videos analysis, thus we utilized the adapted mDISCERN for greater efficiency. However, this choice entails a methodological limitation: other validated instruments such as PEMAT (34), JAMA (22, 23, 32), and HONcode (33) were not employed.

Beyond the choice of instruments, our study expands upon existing procedural paradigms in two key aspects. First, unlike the majority of prior research which typically limits investigation to a single platform (e.g., TikTok only) (21, 23, 26, 27, 35) or dual platforms (21, 36), our study conducts a comprehensive comparative analysis across China's three dominant video platforms (Bilibili, TikTok, and Kwai). This multi-platform approach offers a broader ecological view of the digital information landscape. Second, distinct from the standard practice of relying exclusively on medical professionals for scoring, our experimental design explicitly incorporated non-professional

riters. This inclusion addresses a methodological limitation in previous works by simulating the actual perspective of the lay public, thereby enhancing the ecological validity of the quality assessment.

When comparing our results with these prior frameworks, a distinct “domain-specific risk” becomes evident. Our analysis revealed that dietary weight loss videos generally failed to achieve satisfactory quality scores (median scores = 2; mean scores in the range of 1.7–2.1). These values are systematically lower than those reported in studies of more domain-specific topics on the same platforms. For instance, on TikTok, videos related to thyroid cancer achieved a mean GQS of 3.72 (35) and videos regarding liver cancer reached a median GQS of 3 (21). Notably, this quality gap persists even when compared to dietary guidance for specific clinical conditions. Our scores were lower than those reported for inflammatory bowel disease dietary management on Chinese platforms (mean GQS 2.61, median 3) (19) and post-bariatric surgery nutrition (mean GQS 2.35, median 2) (23). This discrepancy suggests that unlike oncology or surgery content, which retains a higher barrier to entry, the dietary weight loss domain is disproportionately susceptible to low-quality information and lacks the rigor seen in other medical specialties.

4.4 Public health implications

Over the past forty years, more than half of Chinese adult population has become overweight or even obese (37). Obesity can lead to a range of metabolic disorders such as hypertension, hyperglycemia, and hyperuricemia, which pose a serious threat to public health. A multitude of evidence suggests that a weight loss of 5–15% can significantly improve metabolic disorders and reduce the risk of cardiovascular diseases and other weight-related comorbidities (38, 39). As the primary source of health information shifts from medical professionals to the internet, the uneven quality of online content poses new challenges. Low-quality diet videos may spread misinformation and encourage unhealthy behaviors, such as crash dieting or following unbalanced regimens (40, 41).

From a health communication perspective, lifestyle-oriented topics are more conducive to narrative and experience-based communication styles that enhance relatability and audience engagement but may compromise informational rigor. In addition, from the standpoint of digital health literacy frameworks, audiences may possess adequate functional literacy to understand practical dietary advice while lacking the critical literacy required to assess scientific credibility. Collectively, these factors may make the dietary weight loss domain particularly susceptible to the spread of lower-quality information.

Crucially, our study revealed a negative correlation between video quality and engagement metrics. This phenomenon is likely driven by the inherent nature of short-video platforms: users typically consume content during fragmented leisure time, favoring shorter, fast-paced videos. Consequently, algorithmic recommendation mechanisms may be associated with greater visibility for videos with higher completion rates. However, given the cross-sectional nature of our study, these findings should be interpreted as associations only and do not establish a causal relationship between algorithmic recommendation mechanisms and video quality. This pattern may suggest the presence of a “popularity bias,” which may restrict the effective dissemination of evidence-based information to the intended audience. Addressing this potential imbalance may require coordinated efforts at multiple levels.

First, at the policy level, building on the Chinese government’s issuance of the first health promotion guide (42), we recommend that health authorities strengthen the implementation of existing policies by developing clear standards for short-video health content, establishing evaluation and certification mechanisms, and enhancing supervision of misleading information. Second, video platforms should take proactive steps by establishing dedicated medical sections, forming expert review teams to assess medical content, implementing visible credibility indicators (e.g., verified professional badges), and refining video recommendation algorithms to prioritize content quality over mere engagement. Additionally, medical professionals need to improve their digital communication skills and be responsible for the accuracy and reliability of the content they produce. Finally, beyond regulation, there is an urgent need for digital health literacy interventions. Public health education should focus on equipping individuals with critical viewing skills, enabling them to distinguish between scientific advice and misleading narratives in the complex digital environment.

4.5 Limitations and future work

Our study has several limitations. First, the selection of the top 100 videos based on platform algorithms likely introduces an inherent selection bias. While this sampling method effectively captures the information most accessible to the general public, it limits the generalizability of our results to the broader corpus of dietary weight-loss videos. High-quality educational videos that fail to achieve high algorithmic rankings may have been excluded from this analysis. Second, only a single keyword was used to retrieve videos, and the potential use of multiple keywords from a user-centered perspective was not explored. Third, while data were collected within a single day to control for algorithmic volatility, this cross-sectional snapshot may not fully account for temporal fluctuations in ranking and recommendation systems. Fourth, our study focused videos from Chinese platforms and did not assess or compare contents from non-Chinese videos, which limits the generalizability of our findings to international video platforms. Fifth, as a content analysis, this study focuses solely on assessing the quality and reliability of the information provided. We did not investigate the actual impact of these videos on user outcomes, such as behavioral changes or weight loss efficacy. Finally, we acknowledge the statistical issue of multiplicity arising from the numerous comparisons performed. As an exploratory analysis, we did not apply adjustments (e.g., Bonferroni correction) to avoid masking potential signals, inherently inflating the risk of Type 1 errors. Therefore, marginal *p*-values should be interpreted with caution, and these findings warrant validation in future studies with larger sample sizes.

In future work, we aim to adopt a more user-centered approach to simulate real-world information retrieval behaviors. This approach will incorporate multiple relevant keywords (e.g., colloquial terms and synonyms) and adjust sampling strategies to focus on the top-ranked results. In addition, samples will be collected at different time points to account for temporal variations in algorithmic recommendations. Together, these strategies aim to capture the content users are most likely to encounter in real time. Such studies will complement our current landscape analysis and better reflect the diversity of dietary weight-loss information accessible to the public.

5 Conclusion

We collected 100 videos on the theme of dietary weight loss from Bilibili, TikTok and Kwai, and evaluated their reliability and quality using the GQS and mDISCERN tools. After three rounds of scoring, we found that while the GQS and mDISCERN scores for Bilibili were higher than those of TikTok and Kwai, the final scores of the three platforms did not exceed 3 points, which indicated that the video quality of the three video platforms on the theme of dietary weight loss was low. Medical practitioners are encouraged to actively create high-quality medical science popularization videos, providing reliable sources of information for public health education. At the same time, video social media platforms should establish relevant policies to supervise and review the publication of medical science popularization videos, in order to prevent users from receiving incorrect medical knowledge. These combined efforts are important for improving digital health communication and promoting effective dissemination of dietary weight loss information in the digital environment.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

SZ: Formal analysis, Methodology, Software, Writing – original draft. YZ: Data curation, Writing – original draft, Formal analysis, Investigation. XC: Investigation, Validation, Writing – review & editing. QZho: Data curation, Writing – review & editing. FS: Data curation, Writing – review & editing. RX: Data curation, Writing – review & editing. ZS: Data curation, Writing – review & editing. JS: Data curation, Writing – review & editing. LY: Validation, Writing – review & editing. ZX: Data curation, Writing – review & editing. QZha: Data curation, Writing – review & editing. SY: Data curation, Writing – review & editing. YC: Conceptualization, Writing – review & editing, Project administration, Supervision. HY: Conceptualization, Methodology, Writing – review & editing, Project administration. YX: Conceptualization, Resources, Writing – review & editing.

References

- World Health Organization. (2025). Obesity and overweight. Available online at: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> (Accessed March 14, 2026).
- Bray GA. Medical consequences of obesity. *J Clin Endocrinol Metab.* (2004) 89:2583–9. doi: 10.1210/jc.2004-0535
- Singh GM, Danaei G, Farzadfar F, Stevens GA, Woodward M, Wormser D, et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. *PLoS One.* (2013) 8:e65174. doi: 10.1371/journal.pone.0065174
- Czernichow S, Kengne A-P, Stamatakis E, Hamer M, Batty GD. Body mass index, waist circumference, and waist-hip ratio: which is the better discriminator of cardiovascular disease mortality risk? Evidence from an individual-participant meta-analysis of 82,864

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author YX declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2026.1749849/full#supplementary-material>

participants from nine cohort studies. *Obes Rev.* (2011) 12:680–7. doi: 10.1111/j.1467-789X.2011.00879.x

5. Luppino FS, De Wit LM, Bouvy PF, Stijnen T, Cuijpers P, Penninx BWJH, et al. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry.* (2010) 67:220. doi: 10.1001/archgenpsychiatry.2010.2

6. Williams DM, Nawaz A, Evans M. Drug therapy in obesity: a review of current and emerging treatments. *Diabetes Ther.* (2020) 11:1199–216. doi: 10.1007/s13300-020-00816-y

7. Dansinger ML, Gleason JA, Griffith JL, Selker HP, Schaefer EJ. Comparison of the Atkins, Ornish, weight watchers, and zone diets for weight loss and heart disease risk reduction: a randomized trial. *JAMA.* (2005) 293:43–53. doi: 10.1001/jama.293.1.43

8. Alhassan S, Kim S, Bersamin A, King AC, Gardner CD. Dietary adherence and weight loss success among overweight women: results from the a TO Z weight loss study. *Int J Obes.* (2008) 32:985–91. doi: 10.1038/ijo.2008.8
9. Gibson AA, Sainsbury A. Strategies to improve adherence to dietary weight loss interventions in research and real-world settings. *Behav Sci.* (2017) 7:44. doi: 10.3390/bs7030044
10. Bujnowska-Fedak MM. Trends in the use of the internet for health purposes in Poland. *BMC Public Health.* (2015) 15:194. doi: 10.1186/s12889-015-1473-3
11. Kondo Y. Patients' informational needs for pharmacotherapy and health literacy. *Yakugaku Zasshi.* (2021) 141:387–91. doi: 10.1248/yakushi.20-00207-3
12. International Telecommunication Union. Measuring digital development: Facts and figures 2021. (2021). Available online at: <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2021.pdf> (Accessed March 14, 2026).
13. Fox S, Duggan M. 35% of U.S. Adults Have Gone Online to Figure Out a Medical Condition; of these, Half Followed Up with a Visit to a Medical Professional. Pew Research Center (2013).
14. Minadeo M, Pope L. Weight-normative messaging predominates on TikTok—A qualitative content analysis. *PLoS One.* (2022) 17:e0267997. doi: 10.1371/journal.pone.0267997
15. Rus HM, Cameron LD. Health communication in social media: message features predicting user engagement on diabetes-related Facebook pages. *Ann Behav Med.* (2016) 50:678–89. doi: 10.1007/s12160-016-9793-9
16. Feng B, Malloch YZ, Kravitz RL, Verba S, Iosif A-M, Slavik G, et al. Assessing the effectiveness of a narrative-based patient education video for promoting opioid tapering. *Patient Educ Couns.* (2021) 104:329–36. doi: 10.1016/j.pec.2020.08.019
17. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. *Health Informatics J.* (2015) 21:173–94. doi: 10.1177/1460458213512220
18. Sun Y, Zhang Y, Gwizdzka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. *J Med Internet Res.* (2019) 21:e12522. doi: 10.2196/12522
19. He Z, Wang Z, Song Y, Liu Y, Kang L, Fang X, 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. doi: 10.2196/41518
20. Zhang J, Han P, Tang Y, Xi W, Xiao X, Yang F. Popular science and education of cosmetic surgery in China: quality and reliability evaluation of Douyin short videos. *Health Expect.* (2023) 26:1221–6. doi: 10.1111/hex.13737
21. Zheng S, Tong X, Wan D, Hu C, Hu Q, Ke Q. Quality and reliability of liver cancer-related short Chinese videos on TikTok and Bilibili: cross-sectional content analysis study. *J Med Internet Res.* (2023) 25:e47210. doi: 10.2196/47210
22. Sun F, Zheng S, Wu J. Quality of information in gallstone disease videos on TikTok: cross-sectional study. *J Med Internet Res.* (2023) 25:e39162. doi: 10.2196/39162
23. Batar N, Kermen S, Sevdin S, Yildiz N, Güçlü D. Assessment of the quality and reliability of information on nutrition after bariatric surgery on YouTube. *Obes Surg.* (2020) 30:4905–10. doi: 10.1007/s11695-020-05015-z
24. Lahooti A, Hassan A, Critelli B, Westerveld D, Newberry C, Kumar S, et al. Quality and popularity trends of weight loss procedure videos on TikTok. *Obes Surg.* (2023) 33:714–9. doi: 10.1007/s11695-022-06409-x
25. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health.* (1999) 53:105–11. doi: 10.1136/jech.53.2.105
26. Karakoyun A, Yildirim A. YouTube videos as a source of information concerning Behçet's disease: a reliability and quality analysis. *Rheumatol Int.* (2021) 41:2117–23. doi: 10.1007/s00296-021-05009-9
27. Wang L, Li Y, Gu J, Xiao L. A quality analysis of thyroid cancer videos available on TikTok. *Front Public Health.* (2023) 11:9728. doi: 10.3389/fpubh.2023.1049728
28. Singh AG, Singh S, Singh PP. YouTube for information on rheumatoid arthritis—a wakeup call? *J Rheumatol.* (2012) 39:899–903. doi: 10.3899/jrheum.111114
29. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen Van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the world wide web. *Am J Gastroenterol.* (2007) 102:2070–7. doi: 10.1111/j.1572-0241.2007.01325.x
30. Mueller SM, Hongler VNS, Jungo P, Cajacob L, Schwegler S, Steveling EH, 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. doi: 10.2196/15599
31. Azer SA. Are DISCERN and JAMA suitable instruments for assessing YouTube videos on thyroid cancer? Methodological concerns. *J Cancer Educ.* (2020) 35:1267–77. doi: 10.1007/s13187-020-01763-9
32. Ferhatoglu MF, Kartal A, Ekici U, Gurkan A. 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) 29:1477–84. doi: 10.1007/s11695-019-03738-2
33. Xue X, Yang X, Xu W, Liu G, Xie Y, Ji Z. TikTok as an information hodgepodge: evaluation of the quality and reliability of genitourinary cancers related content. *Front Oncol.* (2022) 12:789956. doi: 10.3389/fonc.2022.789956
34. Yang S, Zhan J, Xu X. Is TikTok a high-quality source of information on thyroid cancer? *Endocrine.* (2023) 81:270–6. doi: 10.1007/s12020-023-03332-8
35. Tseng L-W, Lu Y-C, Tseng L-C, Chen Y-C, Chen H-Y. Performance of ChatGPT-4 on Taiwanese traditional Chinese medicine licensing examinations: cross-sectional study. *JMIR Med Educ.* (2025) 11:e58897–7. doi: 10.2196/58897
36. Liu H, Peng J, Li L, Deng A, Huang X, Yin G, 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.* (2024) 11:1296386. doi: 10.3389/fpubh.2023.1296386
37. Zeng Q, Li N, Pan X-F, Chen L, Pan A. Clinical management and treatment of obesity in China. *Lancet Diabetes Endocrinol.* (2021) 9:393–405. doi: 10.1016/S2213-8587(21)00047-4
38. Garvey WT, Mechanick JJ, Brett EM, Garber AJ, Hurley DL, Jastreboff AM, et al. American Association of Clinical Endocrinologists and American College of endocrinology comprehensive clinical practice guidelines for medical Care of Patients with obesity. *Endocr Pract.* (2016) 22:1–203. doi: 10.4158/EP161365.GL
39. Durrer Schutz D, Busetto L, Dicker D, Farpour-Lambert N, Pryke R, Toplak H, et al. European practical and patient-Centred guidelines for adult obesity management in primary care. *Obes Facts.* (2019) 12:40–66. doi: 10.1159/000496183
40. Tabár L, Dean PB, Chen TH-H, Yen AM-F, Chen SL-S, Fann JC-Y, et al. The incidence of fatal breast cancer measures the increased effectiveness of therapy in women participating in mammography screening. *Cancer.* (2019) 125:515–23. doi: 10.1002/cncr.31840
41. Du R-C, Zhang Y, Wang M-H, Lu N-H, Hu Y. TikTok and Bilibili as sources of information on *Helicobacter pylori* in China: a content and quality analysis. *Helicobacter.* (2023) 28:e13007. doi: 10.1111/hel.13007
42. National Health Commission of the People's Republic of China. Guidance on establishing and improving the release and dissemination mechanism of health science knowledge through all media (2022). Available online at: <https://www.nhc.gov.cn/xcs/c100123/202205/373291f407154db388b5383f2eec9bba.shtml> (Accessed March 14, 2026).