

Quality and Content of Abdominal Aortic Aneurysm-Related Short Videos on TikTok, Bilibili, and YouTube: A Cross-Platform Content Analysis

ming wang

Shanghai Ninth People's Hospital

chenghao yang

Zhongshan Hospital

min zhou

Zhongshan Hospital

zhenyu shi

shizhenyu_vascular@163.com

Zhongshan Hospital

Research Article

Keywords: abdominal aortic aneurysm, social media, short-video platforms, health information quality, patient education

Posted Date: July 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10143882/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Cardiovascular Disorders on September 28th, 2026. See the published version at <https://doi.org/10.1186/s12872-026-06709-5>.

Abstract

Background

Short-video platforms are increasingly used by the public to obtain health information; however, the quality of abdominal aortic aneurysm (AAA)-related content across major platforms remains poorly characterized.

Objective

To systematically evaluate the content completeness, educational quality, reliability, and transparency of AAA-related short videos on TikTok, Bilibili, and YouTube.

Methods

We conducted a cross-sectional content analysis. Videos were retrieved on November 19, 2024, using the keywords “abdominal aortic aneurysm” and “腹主动脉瘤.” After screening, 300 videos were included (100 per platform). Video characteristics, uploader type, and presentation format were recorded. Content completeness was assessed across six domains (definition, etiology/risk factors, symptoms, diagnosis, treatment, and prevention) using a 0–2 scale (in 0.5 increments). Video quality and reliability were evaluated using the Global Quality Score (GQS), modified DISCERN, and JAMA benchmark criteria.

Results

YouTube videos were significantly longer (median: 1639 s) than those on TikTok (median: 76 s) and Bilibili (median: 89 s) ($P < 0.001$). Conversely, TikTok demonstrated markedly higher user engagement; the median number of likes was 209 on TikTok, compared with 62 on YouTube and 9 on Bilibili ($P < 0.001$). Overall quality differed significantly across platforms: the median GQS was 3.0 on YouTube, higher than on TikTok (2.0) and Bilibili (2.0), with similar patterns observed for modified DISCERN and JAMA scores (all pairwise $P < 0.001$). Most videos provided incomplete domain coverage, with particularly prominent gaps in risk factors and prevention. Engagement metrics showed weak or no correlations with quality scores, whereas the three quality instruments were strongly intercorrelated ($r = 0.70$ – 0.86 , all $P < 0.001$).

Conclusions

AAA-related short videos vary substantially in quality across platforms. Although YouTube and medically affiliated uploaders tended to provide more reliable content, clinically decision-relevant information remained incomplete across all platforms. Standardized, medically authored content and strengthened

platform governance are warranted to reduce misinformation risk and improve public health education regarding AAA.

Introduction

Abdominal aortic aneurysm (AAA) is a localized dilatation of the abdominal aorta that frequently remains clinically silent until rupture.^{1–3} AAA predominantly affects older adults, particularly men and individuals with a history of smoking, and rupture is associated with catastrophic outcomes; overall mortality has been reported to exceed 80% once rupture occurs⁴. In the United States, AAA rupture accounts for approximately 15,000 deaths annually and remains an important contributor to mortality among older men⁵. This substantial public health burden underscores the critical importance of early detection and timely intervention. Elective repair before rupture markedly improves prognosis, and one-time ultrasound screening in at-risk populations has been shown to reduce AAA-related rupture rates and mortality^{6,7}. However, the effectiveness of screening strategies depends in part on public awareness and risk recognition. As with many “silent” diseases, insufficient knowledge of AAA risk factors, warning signs, and screening recommendations may delay diagnosis and reduce opportunities for preventive care, thereby perpetuating avoidable morbidity and mortality.

In the digital health era, short-video social media platforms have emerged as influential channels for disseminating medical information.^{8,9} Platforms such as TikTok, Bilibili, and YouTube have large global user bases and employ algorithm-driven content personalization, enabling health-related content to reach wide audiences rapidly.¹⁰ Their interactive features—including commenting, sharing, and “likes”—lower barriers to participation and allow both healthcare professionals and lay users to communicate health information beyond traditional clinical and educational settings. At the same time, the decentralized and user-generated nature of social media raises substantial concerns regarding information quality.¹¹ Unlike peer-reviewed publications, these platforms generally lack rigorous quality assurance and standardized editorial oversight, resulting in considerable variability in accuracy, balance, and source transparency. Misinformation and low-quality health advice may consequently spread widely, potentially misleading viewers and contributing to inappropriate health decisions.^{12,13} Prior content analyses across multiple disease areas have consistently reported that a considerable proportion of short videos are educationally suboptimal and of limited reliability, highlighting the need for systematic evaluation of online health information.

Despite increasing scrutiny of health-related content on social media, AAA-related short videos have not been comprehensively assessed across major platforms. This evidence gap is clinically relevant, as unvetted AAA content may undermine effective risk communication for a condition in which timely screening and intervention are critical. Therefore, the present study aimed to systematically evaluate the educational quality, reliability, and transparency of AAA-related short videos on TikTok, Bilibili, and YouTube. Using established assessment instruments—including the Global Quality Score (GQS), modified DISCERN criteria, and Journal of the American Medical Association (JAMA) benchmarks—we

rated each video's overall quality, reliability, and key indicators of source accountability^{14,15}. By identifying strengths and deficiencies in currently available AAA content, this study sought to provide evidence to support higher-quality digital health education and inform platform-level governance strategies aimed at reducing misinformation. Ultimately, improving the quality of AAA-related information accessible to the public may enhance awareness among at-risk individuals, facilitate earlier detection through screening, and mitigate harms associated with inaccurate online health content.

Methods

Study Design and Sampling

We conducted a cross-sectional, cross-platform content analysis of AAA-related short videos on TikTok, Bilibili, and YouTube. Videos were retrieved on November 19, 2024, using platform-specific search queries with the terms “abdominal aortic aneurysm” and “腹主动脉瘤.” The top results returned by each platform in their default ranking order were screened. After removing duplicates, advertisements, and irrelevant entries, 100 eligible videos per platform (total n = 300) were retained for analysis. The complete selection process is illustrated in Fig. 1.

Eligibility Criteria

Videos were included if they primarily addressed AAA-related health information targeting a general audience. Videos were excluded if they were (1) duplicated across searches, (2) primarily promotional or commercial advertisements, (3) unrelated to AAA, or (4) devoid of substantive informational content.

Data Extraction and Video Characteristics

For each included video, we extracted video characteristics and user engagement indicators, including duration (seconds), days since upload, and interaction counts (likes, comments, and—where platform-available—collections/saves and shares). Metrics unavailable for a given platform (e.g., collections and shares on YouTube) were recorded as not available (N/A) by design, rather than treated as zero.

Uploader Classification

Uploaders were categorized into four prespecified groups: (1) medical professionals (e.g., physicians, nurses), (2) medical institutions (e.g., hospitals, medical associations), (3) non-medical professionals (e.g., science communicators, journalists), and (4) non-medical institutions (e.g., commercial health channels, general media outlets). Platform-level distributions were summarized to characterize the creator ecosystem on each platform.

Presentation Format and Content Completeness

Presentation format was coded into five predefined categories: monologue, dialogue, animation, vlog, and news report. To evaluate informational completeness, video content was independently assessed across six domains: (1) definition, (2) etiology/risk factors, (3) symptoms, (4) diagnosis, (5) treatment, and (6) prevention. Each domain was rated on a 0–2 scale (in 0.5 increments; i.e., 0, 0.5, 1.0, 1.5, or 2.0), with higher scores indicating more complete coverage of the corresponding topic.

Quality and Reliability Assessment

Overall informational quality and reliability were assessed using three validated instruments: (1) the Global Quality Score (GQS), a five-point scale evaluating educational value and flow; (2) the modified DISCERN tool, which assesses reliability and treatment-related information quality; and (3) JAMA benchmark criteria, which evaluate authorship, attribution, currency, and disclosure. (detailed scoring criteria are provided in Supplementary Files 1–3) Scores were computed for each video and compared across platforms and uploader categories.

Statistical Analysis

Given the ordinal nature of most outcomes and the non-normal distribution of engagement metrics, continuous and ordinal variables are reported as median (interquartile range [IQR]), and categorical variables as n (%). Group differences across the three platforms were tested using the Kruskal-Wallis test for continuous/ordinal variables and χ^2 -based methods for categorical variables, with post hoc pairwise comparisons applied when overall tests were statistically significant. Spearman rank correlation coefficients (r) and two-sided P values were used to examine associations between engagement indicators and quality/reliability scores; only associations with $P < 0.05$ were interpreted as statistically significant. In figures, * denotes $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

Ethics Considerations

All data were derived from publicly available online videos and did not involve interaction with human participants or collection of identifiable private information. Formal ethics approval was therefore not required.

Results

Basic Characteristics of Included Videos

A total of 300 AAA-related videos were included after the removal of duplicates, advertisements, and irrelevant content, as illustrated in the screening flowchart (Fig. 1). The basic characteristics of the included videos are presented in Table 1.

Table 1
Baseline characteristics of included videos across TikTok, Bilibili, and YouTube.

Variables	Total (n = 300)	TikTok (n = 100)	Bilibili (n = 100)	YouTube (n = 100)	P
Video duration, s, median (IQR)	129.00 (61.50, 1152.75)	76.00 (36.75, 98.50)	89.00 (50.75, 177.25)	1639.00 (907.75, 3070.50)	< 0.001
Days since upload, median (IQR)	361.00 (138.75, 757.25)	344.50 (142.50, 743.50)	614.50 (344.50, 957.25)	175.50 (102.00, 485.00)	< 0.001
Likes, median (IQR)	59.50 (9.00, 387.50)	209.00 (65.25, 909.00)	9.00 (2.00, 53.25)	62.00 (13.00, 488.50)	< 0.001
Collections, median (IQR)	16.00 (3.00, 65.00)	28.50 (7.00, 101.00)	7.00 (1.00, 44.00)	N/A	< 0.001
Comments, median (IQR)	3.00 (1.00, 31.00)	25.00 (5.00, 93.50)	1.00 (1.00, 5.00)	1.00 (0.00, 23.00)	< 0.001
Shares, median (IQR)	10.00 (2.00, 42.50)	17.50 (4.00, 79.00)	5.00 (0.50, 24.00)	N/A	< 0.001
GQS, median (IQR)	2.00 (2.00, 3.00)	2.00 (2.00, 3.00)	2.00 (1.00, 2.00)	3.00 (2.00, 4.00)	< 0.001
Modified DISCERN, median (IQR)	2.00 (1.00, 3.00)	2.00 (1.00, 2.00)	1.00 (1.00, 2.00)	3.00 (2.00, 3.00)	< 0.001
JAMA benchmark score, median (IQR)	2.00 (1.00, 3.00)	2.00 (1.00, 2.25)	1.00 (1.00, 1.00)	3.00 (2.00, 3.00)	< 0.001
<i>Abbreviations: IQR, interquartile range; GQS, Global Quality Score; N/A, not available (metric not provided by platform).</i>					

Video duration differed markedly across platforms. TikTok videos had a median duration of 76.00 s (IQR 36.75–98.50), Bilibili videos 89.00 s (IQR 50.75–177.25), and YouTube videos 1639.00 s (IQR 907.75–3070.50) ($P < 0.001$). The time since upload also varied significantly: Bilibili videos had the longest median of 614.50 days (IQR 344.50–957.25), followed by TikTok at 344.50 days (IQR 142.50–743.50) and YouTube at 175.50 days (IQR 102.00–485.00) ($P < 0.001$).

User engagement metrics showed considerable variation across platforms. TikTok videos received a median of 209.00 likes (IQR 65.25–909.00), substantially higher than Bilibili (9.00; IQR 2.00–53.25) and YouTube (62.00; IQR 13.00–488.50) ($P < 0.001$). For collections, TikTok showed a median of 28.50 (IQR 7.00–101.00) and Bilibili 7.00 (IQR 1.00–44.00); this metric was unavailable for YouTube. Shares followed a similar pattern, with TikTok at 17.50 (IQR 4.00–79.00) and Bilibili at 5.00 (IQR 0.50–24.00); data were unavailable for YouTube. Comments differed significantly, with TikTok videos receiving a median of 25.00 (IQR 5.00–93.50) compared with 1.00 for both YouTube (IQR 0.00–23.00) and Bilibili

(IQR 1.00–5.00) ($P < 0.001$). Significant differences were observed across the three platforms for all reported video characteristics (all $P < 0.001$).

Distribution of Uploaders

The distribution of uploaders across the included videos is shown in Fig. 2. Among all 300 videos, medical professionals constituted the largest contributor group ($n = 130$), followed by medical institutions ($n = 87$), non-medical professionals ($n = 48$), and non-medical institutions ($n = 35$).

Platform-specific distributions revealed distinct patterns. TikTok included 60 videos from medical professionals, 19 from medical institutions, 12 from non-medical institutions, and 9 from non-medical professionals. Bilibili contained 49 videos from medical professionals, 23 from non-medical professionals, 18 from medical institutions, and 10 from non-medical institutions. YouTube differed markedly, with 50 videos from medical institutions, 21 from medical professionals, 16 from non-medical professionals, and 13 from non-medical institutions. These results indicate that although medically affiliated sources were the primary contributors overall, the distribution of uploader types varied substantially across platforms.

Content Characteristics

The distribution of presentation formats is shown in Fig. 3 and Supplementary Table S1. Among all 300 videos, monologue was the most common format ($n = 100$, 33.33%), followed by dialogue ($n = 77$, 25.67%), animation ($n = 66$, 22.00%), vlog ($n = 39$, 13.00%), and news report ($n = 18$, 6.00%). Significant platform-level differences were observed ($P < 0.001$): Bilibili had a higher proportion of animation videos (31.00%), TikTok was dominated by dialogue format (40.00%), and YouTube showed a greater proportion of monologue presentations (38.00%).

Content completeness scores across the six domains are summarized in Table 2. For definition, 43.00% of videos scored 0.5, with 21.00% scoring 0 ($P < 0.001$). For etiology/risk factors, 45.67% scored 0 and 25.67% scored 0.5, with no significant platform-level differences ($P = 0.173$). For symptoms, 40.33% scored 0 and 33.00% scored 0.5 ($P = 0.011$). For diagnosis, 39.00% scored 0 and 34.33% scored 0.5 ($P = 0.044$). For treatment, 31.00% scored 0.5 and 27.33% scored 1 ($P < 0.001$). For prevention, more than half of the videos (51.67%) scored 0, followed by 29.00% scoring 0.5 ($P = 0.036$). Overall, most videos provided only partial domain coverage (scores of 0.5–1.0), while complete coverage (score of 2.0) was uncommon across all domains.

Table 2
Distribution of content completeness scores across the three platforms.

Domain / Score	Total (n = 300) n (%)	TikTok (n = 100) n (%)	Bilibili (n = 100) n (%)	YouTube (n = 100) n (%)	P
Definition					< 0.001
0	63 (21.00)	35 (35.00)	10 (10.00)	18 (18.00)	
0.5	129 (43.00)	38 (38.00)	56 (56.00)	35 (35.00)	
1	52 (17.33)	18 (18.00)	15 (15.00)	19 (19.00)	
1.5	46 (15.33)	6 (6.00)	17 (17.00)	23 (23.00)	
2	10 (3.33)	3 (3.00)	2 (2.00)	5 (5.00)	
Etiology/Risk Factors					0.173
0	137 (45.67)	52 (52.00)	42 (42.00)	43 (43.00)	
0.5	77 (25.67)	26 (26.00)	30 (30.00)	21 (21.00)	
1	41 (13.67)	7 (7.00)	15 (15.00)	19 (19.00)	
1.5	35 (11.67)	13 (13.00)	8 (8.00)	14 (14.00)	
2	10 (3.33)	2 (2.00)	5 (5.00)	3 (3.00)	
Symptoms					0.011
0	121 (40.33)	31 (31.00)	37 (37.00)	53 (53.00)	
0.5	99 (33.00)	46 (46.00)	34 (34.00)	19 (19.00)	
1	36 (12.00)	10 (10.00)	13 (13.00)	13 (13.00)	
1.5	30 (10.00)	8 (8.00)	10 (10.00)	12 (12.00)	
2	14 (4.67)	5 (5.00)	6 (6.00)	3 (3.00)	
Diagnosis					0.044
0	117 (39.00)	34 (34.00)	37 (37.00)	46 (46.00)	
0.5	103 (34.33)	43 (43.00)	40 (40.00)	20 (20.00)	
1	31 (10.33)	8 (8.00)	11 (11.00)	12 (12.00)	
1.5	41 (13.67)	13 (13.00)	9 (9.00)	19 (19.00)	
2	8 (2.67)	2 (2.00)	3 (3.00)	3 (3.00)	
Treatment					< 0.001

Domain / Score	Total (n = 300) n (%)	TikTok (n = 100) n (%)	Bilibili (n = 100) n (%)	YouTube (n = 100) n (%)	P
Definition					< 0.001
0	68 (22.67)	10 (10.00)	21 (21.00)	37 (37.00)	
0.5	93 (31.00)	46 (46.00)	30 (30.00)	17 (17.00)	
1	82 (27.33)	33 (33.00)	25 (25.00)	24 (24.00)	
1.5	48 (16.00)	10 (10.00)	19 (19.00)	19 (19.00)	
2	9 (3.00)	1 (1.00)	5 (5.00)	3 (3.00)	
Prevention					0.036
0	155 (51.67)	57 (57.00)	54 (54.00)	44 (44.00)	
0.5	87 (29.00)	31 (31.00)	28 (28.00)	28 (28.00)	
1	25 (8.33)	6 (6.00)	4 (4.00)	15 (15.00)	
1.5	22 (7.33)	2 (2.00)	11 (11.00)	9 (9.00)	
2	11 (3.67)	4 (4.00)	3 (3.00)	4 (4.00)	
<i>Scores range from 0 to 2 in 0.5 increments; higher scores indicate more complete coverage of the corresponding domain. P values are from Kruskal-Wallis tests across the three platforms.</i>					

Quality and Reliability Assessment

Quality and reliability scores are summarized in Table 1 and illustrated in Fig. 4. Significant platform-level differences were observed for all three scoring instruments (all $P < 0.001$). For GQS, median scores were 3.00 (IQR 2.00–4.00) for YouTube, 2.00 (IQR 2.00–3.00) for TikTok, and 2.00 (IQR 1.00–2.00) for Bilibili. YouTube scored significantly higher than both other platforms, and TikTok scored significantly higher than Bilibili (all pairwise $P < 0.001$). Similar patterns were observed for modified DISCERN (YouTube: 3.00, IQR 2.00–3.00; TikTok: 2.00, IQR 1.00–2.00; Bilibili: 1.00, IQR 1.00–2.00) and JAMA scores (YouTube: 3.00, IQR 2.00–3.00; TikTok: 2.00, IQR 1.00–2.25; Bilibili: 1.00, IQR 1.00–1.00).

Quality and reliability scores also differed significantly by uploader type (Supplementary Table S2). Videos from medical institutions and medical professionals consistently achieved higher GQS, modified DISCERN, and JAMA scores, whereas content from non-medical institutions and non-medical professionals showed uniformly lower scores (most pairwise comparisons $P < 0.001$). This pattern was consistent across all three scoring instruments.

Correlation Analysis

Correlation results are shown in Fig. 5. The strongest associations were observed among the three quality assessment tools: GQS–modified DISCERN ($r = 0.78$, $P < 0.001$), GQS–JAMA ($r = 0.70$, $P < 0.001$),

and modified DISCERN–JAMA ($r = 0.86$, $P < 0.001$). Substantial intercorrelations were also identified among user engagement indicators: likes correlated with collections ($r = 0.35$, $P < 0.001$), comments ($r = 0.78$, $P < 0.001$), and shares ($r = 0.42$, $P < 0.001$); collections with comments ($r = 0.64$, $P < 0.001$) and shares ($r = 0.79$, $P < 0.001$); and comments with shares ($r = 0.66$, $P < 0.001$).

Several engagement indicators demonstrated weak but statistically significant associations with quality measures. Collections correlated positively with modified DISCERN ($r = 0.20$, $P < 0.01$) and JAMA ($r = 0.20$, $P < 0.01$). Shares similarly correlated with modified DISCERN ($r = 0.18$, $P < 0.01$) and JAMA ($r = 0.20$, $P < 0.01$). By contrast, likes, comments, and days since upload showed no statistically significant associations with any of the three quality instruments (all $P > 0.05$). Overall, the three quality tools were strongly intercorrelated, engagement indicators formed a distinct cluster, and only collections and shares demonstrated weak positive associations with quality scores.

Discussion

Principal Findings

In this cross-platform evaluation of AAA-related short videos, we identified substantial differences in video characteristics, uploader composition, information completeness, and overall quality across TikTok, Bilibili, and YouTube. YouTube videos were generally longer, whereas TikTok videos were shorter but garnered higher engagement metrics.¹⁶ These patterns likely reflect platform-specific content ecosystems and user consumption behaviors: longer-form content may facilitate more structured explanations, while short-form videos are optimized for rapid viewing and audience interaction.

Uploader composition differed across platforms, although medically affiliated sources constituted the primary contributor group on all three platforms. On TikTok, AAA videos were predominantly uploaded by medical professionals, followed by medical institutions. On Bilibili, medical professionals remained the largest contributor group, with a comparatively higher proportion of non-medical professionals. In contrast, YouTube content was dominated by medical institutions, with medical professionals as the second largest contributor group.¹⁷

Consistent with this compositional difference, YouTube demonstrated significantly higher quality and reliability scores across GQS, modified DISCERN, and JAMA benchmarks, and videos from medically affiliated uploaders consistently achieved higher scores than non-medical sources.¹⁸ Nevertheless, none of the platforms achieved uniformly high performance across all metrics, indicating persistent opportunities for improvement in online AAA education.

Correlation analysis confirmed strong agreement among the three quality instruments (GQS–modified DISCERN: $r = 0.78$; GQS–JAMA: $r = 0.70$; modified DISCERN–JAMA: $r = 0.86$; all $P < 0.001$), supporting their convergent validity for assessing online health video quality. In contrast, engagement indicators were not reliable proxies for quality: likes and comments were not significantly associated with quality or

reliability scores ($P > 0.05$), while collections/saves and shares showed only weak positive correlations with modified DISCERN and JAMA ($r \approx 0.18\text{--}0.20$, $P < 0.01$). Days since upload was not significantly associated with quality scores.

Platform Differences in Content and Sources

The observed cross-platform differences likely stem from distinct platform designs, recommendation mechanisms, and creator ecosystems. TikTok's short-video format encourages concise delivery and fast-paced narrative structures, which may increase engagement but also constrain the depth of medical explanations.¹⁹ Bilibili, with its mixed creator community, facilitates diverse content styles but introduces variability in expertise levels and content completeness. YouTube's longer-form format and higher representation of medical institutions may support more structured content delivery and inclusion of key educational elements, which may partially account for its superior quality and reliability performance in the present study.

In addition to uploader identity, presentation format shapes how medical concepts are communicated. Across platforms, monologue, dialogue, and animation were the predominant formats. These approaches may improve accessibility when designed appropriately; however, format alone does not guarantee content completeness or factual accuracy. In the context of AAA education, the critical issue is whether videos provide a minimum "decision-relevant" knowledge set—particularly regarding screening indications, diagnostic pathways, risk communication, and treatment thresholds.²⁰

Content Completeness: Practical Implications and Common Misconceptions

A key concern highlighted by our findings is the limited completeness across core AAA knowledge domains, particularly on short-form platforms.²¹ Across all six assessed dimensions—definition, etiology/risk factors, symptoms, diagnosis, treatment, and prevention—low or partial coverage predominated. Notably, nearly half of all videos scored 0 on etiology/risk factors (45.67%), and more than half scored 0 on prevention (51.67%). Similar patterns were observed for symptoms (40.33% scored 0) and diagnosis (39.00% scored 0), indicating frequent omission of essential educational content.

Even in domains with relatively better coverage, partial scores (0.5–1.0) were far more common than complete coverage, suggesting that users are more likely to receive fragmented messages than a coherent understanding of the disease. From a public health perspective, these content gaps are clinically meaningful because they may reinforce common patient misconceptions and delay appropriate care-seeking or preventive screening. Based on the observed gaps and recurring themes in AAA health communication, several misconceptions warrant explicit clarification in future educational content:

(1) Definition

AAA is sometimes conflated with "tumor/cancer," or small aneurysms are portrayed as harmless and self-resolving, which may undermine adherence to surveillance and follow-up.

(2) Risk factors

Viewers may believe that smoking cessation years prior eliminates residual risk, or that well-controlled blood pressure alone fully prevents AAA, potentially reducing screening uptake among high-risk individuals.

(3) Symptoms

The absence of abdominal pain is frequently interpreted as “no risk,” while normal abdominal pulsation may be misinterpreted as imminent rupture, engendering either false reassurance or unnecessary alarm.

(4) Diagnosis

A misconception persists that physical palpation can definitively rule AAA in or out²². Conversely, exaggerated fear of imaging—particularly computed tomography angiography (CTA)—may lead patients to avoid necessary evaluation.

(5) Treatment: Two opposing extremes commonly coexist: “diagnosis implies immediate surgery” versus “EVAR is a permanent one-time fix.” Open surgical repair may also be framed as categorically inappropriate rather than a valid, context-dependent option.²³

(6) Prevention

Some individuals assume that routine health check-ups automatically include AAA screening, or that lifestyle modification alone can cause an aneurysm to regress.

These misconceptions underscore why our findings carry implications beyond documenting platform-level differences. In a short-video ecosystem, high engagement does not equate to high educational quality, and the most decision-relevant information—screening eligibility, diagnostic standards and the respective roles of ultrasound and CTA, size-based surveillance and intervention thresholds, and post-treatment follow-up protocols—is particularly vulnerable to dilution by time constraints and fragmented storytelling. Future AAA health education strategies should therefore prioritize delivery of a minimum closed-loop message set: (1) who should be screened and when; (2) what constitutes definitive diagnostic evaluation and what each test is used for; (3) what size and clinical thresholds guide surveillance versus intervention; and (4) how follow-up is structured after diagnosis or repair. These messages should preferentially be delivered through medically affiliated accounts, using standardized, reusable scripts and visual aids to reduce the propagation of misconceptions.

User Engagement Versus Information Quality

Our correlation results further confirm that engagement metrics should not be used as surrogates for information quality.²⁴ While engagement indicators were strongly intercorrelated (reflecting popularity dynamics), neither likes nor comments demonstrated statistically significant associations with quality or reliability scores ($P > 0.05$). Only collections/saves and shares showed weak positive associations with

modified DISCERN and JAMA ($r \approx 0.18-0.20$, $P < 0.01$), suggesting that users' deeper interaction behaviors may be marginally better aligned with content reliability than superficial engagement; however, the effect size remains small and insufficient to serve as a quality filter. Platform-level curation mechanisms and visible credibility labeling may therefore be required to help users identify higher-quality AAA information.

Implications for Health Communication

Our findings suggest several practical directions for improving AAA communication in the short-video ecosystem. First, these platforms should be treated as essential but currently suboptimal channels for AAA education; content must be systematically optimized for accuracy and completeness. Second, increasing the supply of medically authored videos is likely to improve overall quality, although production must be adapted to short-video format constraints through structured micro-modules, clear risk communication messaging, and evidence-based visual explanations. Third, platforms could prioritize credible sources within recommendation algorithms and implement visible credibility indicators for viewers. Finally, structured collaborations between clinical specialists and experienced content creators may help reconcile medical accuracy with platform-native storytelling conventions.

Limitations

Several limitations should be acknowledged. First, we sampled a fixed number of videos per platform, which may not represent the full content landscape and may be influenced by search ranking algorithms; results may therefore vary with different search parameters or time points. Second, engagement metrics are time-sensitive and may have changed after data collection. Third, quality scoring instruments, although widely validated, involve some degree of subjective judgment and may not capture all disease-specific inaccuracies. Finally, our analysis reflects the search language and platform context employed, limiting generalizability to other linguistic or cultural settings.

Conclusions

Across TikTok, Bilibili, and YouTube, AAA-related short videos exhibited significant cross-platform differences in video characteristics, uploader composition, content completeness, and information quality. Overall quality and reliability were highest on YouTube, and videos from medically affiliated uploaders consistently outperformed non-medical sources across all three assessment instruments. However, content coverage across core AAA domains remained largely incomplete, with full domain coverage being uncommon, indicating substantial room for improvement in online AAA education. Correlation analyses demonstrated strong convergent validity among GQS, modified DISCERN, and JAMA, while engagement metrics were poor proxies for quality: likes and comments were not significantly associated with quality scores, and only collections and shares showed weak positive associations. These findings support the need for standardized, clinically decision-relevant AAA educational content delivered by credible medical sources, alongside platform-level governance mechanisms to promote accurate public health information.²⁵

Declarations

Ethics Approval and Consent to Participate

Not applicable. This study analyzed only publicly available online content and did not involve human participants.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

Funding

This research did not receive any specific funding from public, commercial, or not-for-profit funding agencies.

Author Contribution

M.W. and C.Y. contributed equally to this work. M.W. and C.Y. conceived and designed the study, performed video retrieval, screening, coding, and data extraction, conducted the statistical analyses, prepared the tables and figures, and drafted the original manuscript. M.Z. and Z.S. provided methodological guidance, supervised the study, and critically revised the manuscript for important intellectual content. All authors contributed to data interpretation, reviewed and approved the final version of the manuscript, agreed to its submission, and agree to be accountable for all aspects of the work.

Acknowledgments

Not applicable.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used ChatGPT (OpenAI) to assist with language editing and clarity improvement. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for all content of the publication.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding authors on reasonable request.

References

1. Golledge J, Thanigaimani S, Powell JT, Tsao PS. Pathogenesis and management of abdominal aortic aneurysm. *Eur Heart J*. 2023;44(29):2682–2697. doi:10.1093/eurheartj/ehad386
2. Puertas-Umbert L, Almendra-Pegueros R, Jiménez-Altayó F, et al. Novel pharmacological approaches in abdominal aortic aneurysm. *Clin Sci (Lond)*. 2023;137(15):1167–1194. doi:10.1042/CS20220795
3. Zhu JX, Tang QQ, Zhou C, Shi XC, Yi SY, Yang Y. Establishment of a New Abdominal Aortic Aneurysm Model in Rats by a Retroperitoneal Approach. *Front Cardiovasc Med*. 2022;9:808732. doi:10.3389/fcvm.2022.808732
4. Goyal A, Saeed H, Shahnoor S, et al. Mortality trends, sex, and racial disparities in older adults due to abdominal aortic aneurysm: a nationwide cross-sectional analysis. *Int J Surg*. 2024;110(12):8241–8245. doi:10.1097/JS9.0000000000002114
5. Lu H, Rateri DL, Bruemmer D, Cassis LA, Daugherty A. Novel mechanisms of abdominal aortic aneurysms. *Curr Atheroscler Rep*. 2012;14(5):402–412. doi:10.1007/s11883-012-0271-y
6. Vats S, Sundquist K, Grundberg A, et al. Hemoglobin homeostasis in abdominal aortic aneurysm: diagnostic and prognostic potential of hemoglobin/heme and scavenger molecules. *BMC Cardiovasc Disord*. 2024;24(1):452. doi:10.1186/s12872-024-04131-3
7. Han Q, Qiao L, Yin L, Sui X, Shao W, Wang Q. The effect of exercise training intervention for patients with abdominal aortic aneurysm on cardiovascular and cardiorespiratory variables: an updated meta-analysis of randomized controlled trials. *BMC Cardiovasc Disord*. 2024;24(1):80. doi:10.1186/s12872-024-03745-x
8. Xu J, Gao Y, Xu X, Li J, Shan Z, Xiao L. Blood or buzz? Decoding the quality of CBC-related short videos on Chinese platforms. *Digit Health*. 2025;11:20552076251375721. doi:10.1177/20552076251375721
9. Yi S, Guo Y, Lin Z, Cao C. Is information evaluated subjectively? Social media has changed the way users search for medical information. *Digit Health*. 2024;10:20552076241259039. doi:10.1177/20552076241259039
10. Zhou Y, Zeng X, Yuan T, et al. Content accuracy and reliability of pulmonary nodule information on social media platforms: a cross-platform study of YouTube, Bilibili, and TikTok. *Front Med (Lausanne)*. 2025;12:1613526. doi:10.3389/fmed.2025.1613526
11. Ghalavand H, Nabilahi A. Exploring online health information quality criteria on social media: a mixed method approach. *BMC Health Serv Res*. 2024;24(1):1311. doi:10.1186/s12913-024-11838-8

12. Helpful or harmful? Navigating the impact of social media influencers' health advice: insights from health expert content creators - PubMed. Accessed December 3, 2025.
<https://pubmed.ncbi.nlm.nih.gov/39696170/>
13. Rivera YM, Corpuz K, Karver TS. Engagement With and Use of Health Information on Social Media Among US Latino Individuals: National Cross-Sectional Survey Study. *J Med Internet Res*. 2025;27:e59387. doi:10.2196/59387
14. Ozduran E, Büyükçoban S. Evaluating the readability, quality and reliability of online patient education materials on post-covid pain. *PeerJ*. 2022;10:e13686. doi:10.7717/peerj.13686
15. Stelmach J, Rychlik J, Zawadzka M, Mazurkiewicz-Beldzinska M. YouTube as a Source of Patient Information for Cerebral Palsy. *Healthcare (Basel)*. 2025;13(13):1492. doi:10.3390/healthcare13131492
16. Javidan A, Vignarajah M, Nelms MW, et al. YouTube as a Source of Patient and Trainee Education in Vascular Surgery: A Systematic Review. *EJVES Vasc Forum*. 2024;61:62–76. doi:10.1016/j.ejvsf.2024.01.054
17. Luo C, Bian L, Jiang L, Liang W, Wu Z. Does YouTube Provide Qualified Patient Education Videos About Atrial Fibrillation? *Front Public Health*. 2022;10:925691. doi:10.3389/fpubh.2022.925691
18. Nie SJ, Ning SN, Ding QC, Hou J, Luo YY, Wang HY. Quality Analysis of Stroke-Related Videos on Video Platforms: Cross-Sectional Study. *JMIR Form Res*. 2025;9:e80458. doi:10.2196/80458
19. Chen Y, Wang Q, Huang X, et al. The quality and reliability of short videos about thyroid nodules on Bilibili and TikTok: Cross-sectional study. *Digit Health*. 2024;10:20552076241288831. doi:10.1177/20552076241288831
20. Liang M, Yang F, Lu S, Zhu C, Wang J, Yan W. Quality and reliability of prostate cancer-Videos on TikTok and Bilibili: Cross-sectional content analysis study. *Digit Health*. 2025;11:20552076251376263. doi:10.1177/20552076251376263
21. Karagoz B, Bakir M, Kececi T. Evaluation of the Accuracy and Quality of Information in Videos About Lateral Epicondylitis Shared on Internet Video Sharing Services. *Cureus*. 2022;14(2):e22583. doi:10.7759/cureus.22583
22. Chiu IM, Chen TY, Zheng YC, et al. Prospective clinical evaluation of deep learning for ultrasonographic screening of abdominal aortic aneurysms. *NPJ Digit Med*. 2024;7(1):282. doi:10.1038/s41746-024-01269-4
23. Kargul M, Skórka P, Gutowski P, Kazimierczak A, Rynio P. Empowering EVAR: Revolutionizing Patient Understanding and Qualification with 3D Printing. *J Cardiovasc Dev Dis*. 2024;11(11):365. doi:10.3390/jcdd11110365
24. Gurler D, Buyukceran I. Assessment of the Medical Reliability of Videos on Social Media: Detailed Analysis of the Quality and Usability of Four Social Media Platforms (Facebook, Instagram, Twitter, and YouTube). *Healthcare (Basel)*. 2022;10(10):1836. doi:10.3390/healthcare10101836
25. Altun A, Askin A, Sengul I, Aghazada N, Aydin Y. Evaluation of YouTube videos as sources of information about complex regional pain syndrome. *Korean J Pain*. 2022;35(3):319–326.

Figures

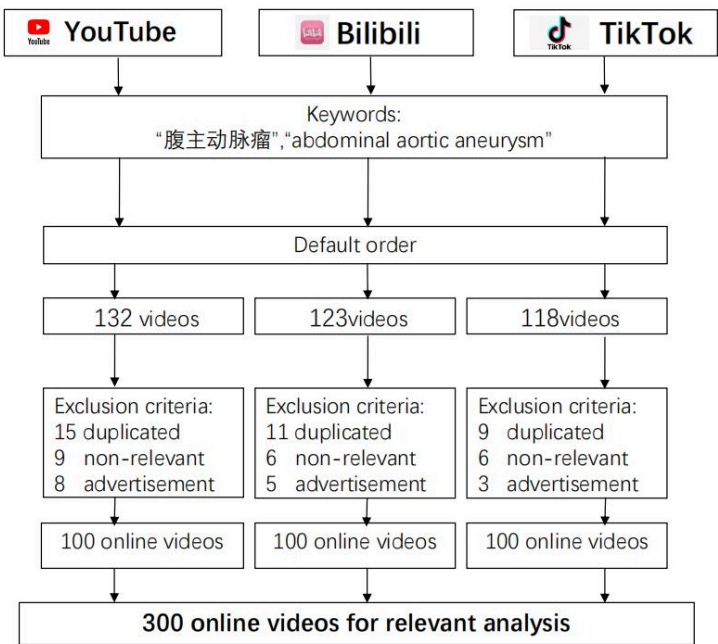


Figure 1

Flowchart of the video selection process.

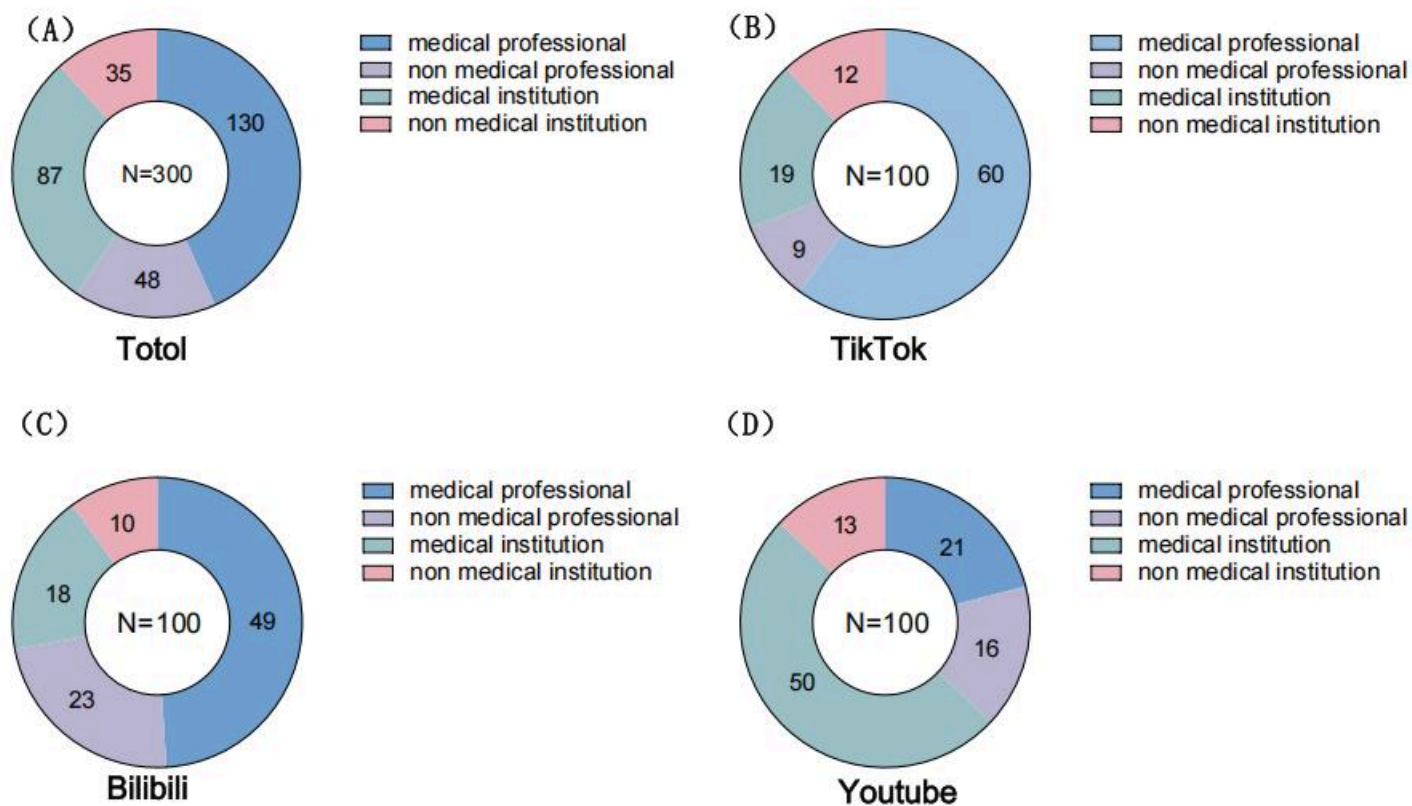


Figure 2

Distribution of uploader types across the three platforms. (A) Overall distribution across all 300 included videos. (B) Distribution among TikTok videos (n = 100). (C) Distribution among Bilibili videos (n = 100). (D) Distribution among YouTube videos (n = 100).

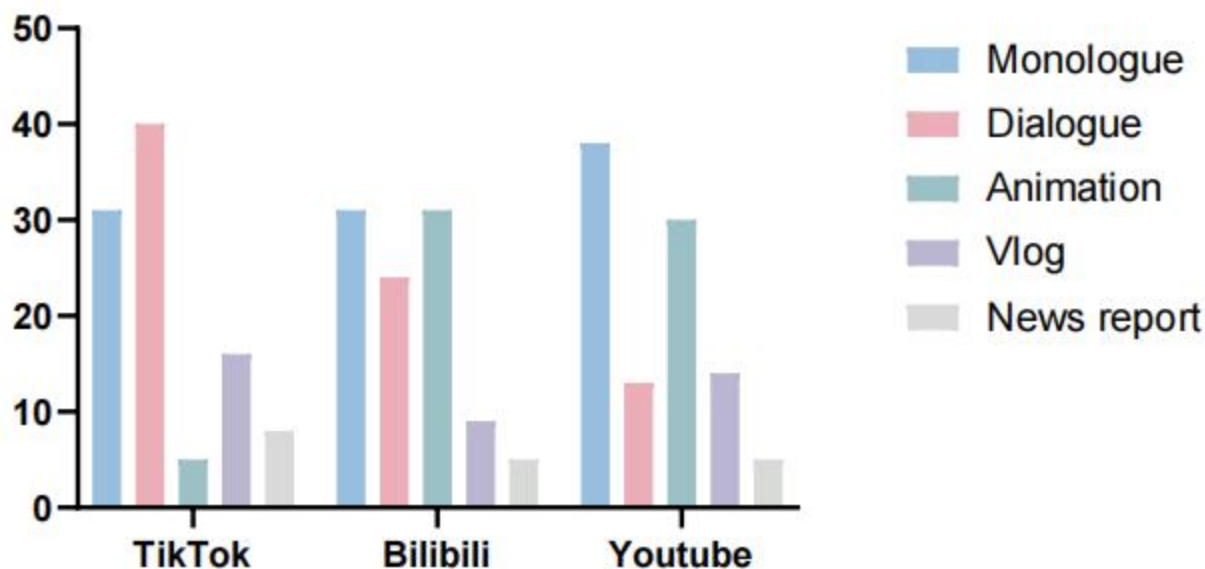


Figure 3

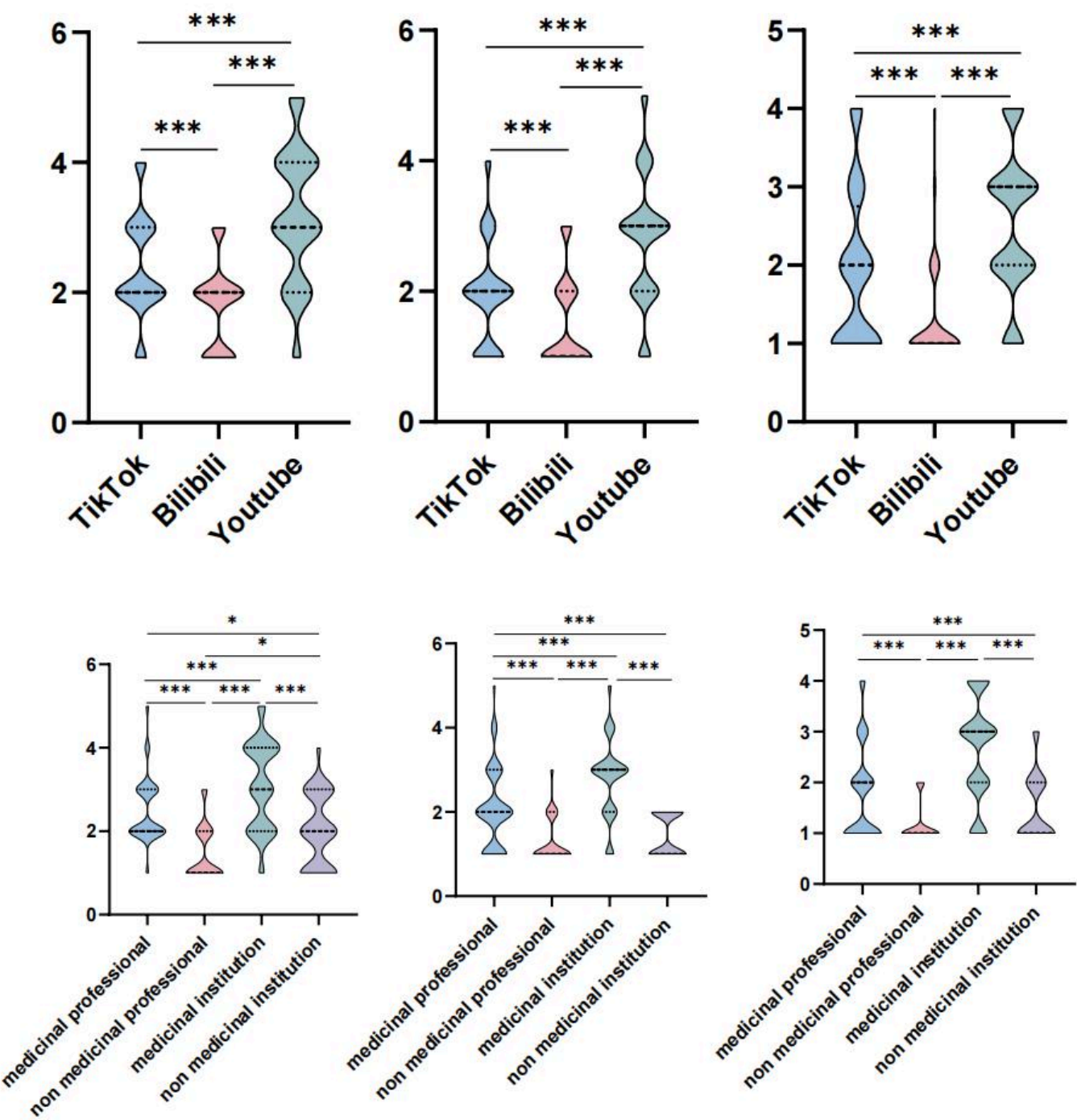


Figure 4

Comparison of quality and reliability scores across platforms and uploader types. (A) GQS scores across the three platforms. (B) Modified DISCERN scores across the three platforms. (C) JAMA benchmark scores across the three platforms. (D) GQS scores stratified by uploader category (medical

professionals, non-medical professionals, medical institutions, non-medical institutions). (E) Modified DISCERN scores by uploader category. (F) JAMA benchmark scores by uploader category.

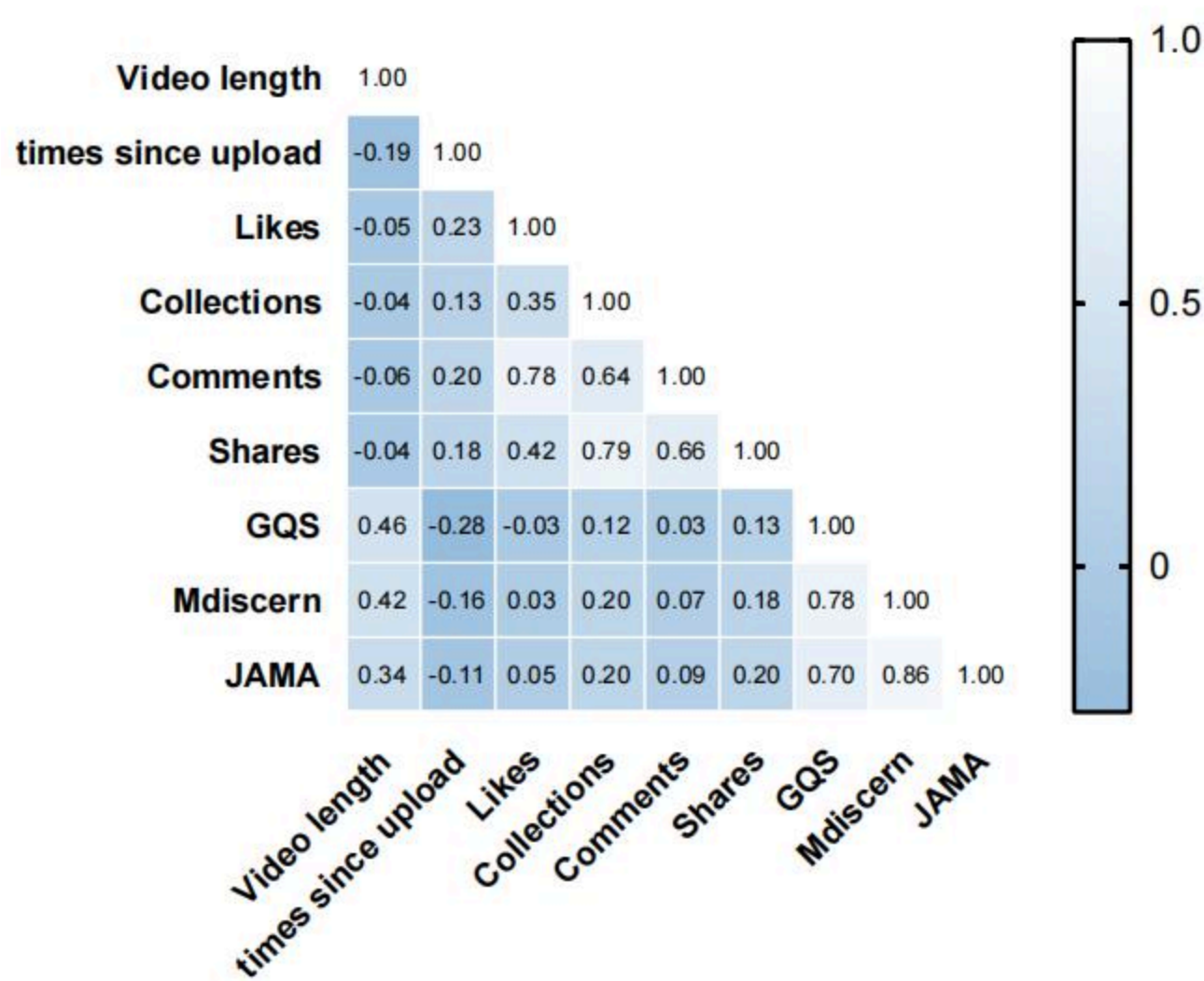


Figure 5

Spearman correlation matrix for video characteristics, engagement indicators, and quality scores. Color intensity reflects the magnitude of the correlation coefficient (r). Only statistically significant correlations (P < 0.05) are interpreted.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementarymaterial1.docx](#)
- [supplementarymaterials2.docx](#)
- [supplementarymaterials3.docx](#)