Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Quality, reliability, and educational value of adjustment-disorder short videos on Chinese social media platforms: a cross-platform content analysis

You Jiang
Zhejiang Chinese Medical University
Chaoyu He
Zhejiang Chinese Medical University
Xuanyi Yu
Zhejiang Chinese Medical University
Jiani Ji
Zhejiang Chinese Medical University
Hongli Yang
Zhejiang Chinese Medical University
Han Yan
Zhejiang Chinese Medical University
Liandong Chen
Zhejiang Chinese Medical University
Xiang Chen
Zhejiang Chinese Medical University
Liangliang Chen
20181055@zcmu.edu.cn

Zhejiang Chinese Medical University

Research Article

Keywords: Adjustment disorder, Short-video platforms, Digital health, Information quality, Social media

Posted Date: May 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9462694/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Widely prevalent worldwide and increasingly noted in China, adjustment disorder (AD) is frequently featured on short-video apps; however, because platforms like TikTok, Bilibili, and RedNote lack robust regulation of health content, the potential for inaccurate information is high, motivating our systematic comparison of the quality of AD-related videos across these three platforms.

Methods

We conducted a cross-sectional content analysis of publicly available Mandarin videos retrieved with the keyword "□□□□" (adjustment disorder). From the top 100 search results on each platform, 228 unique videos were included (Tiktok 78; Bilibili 62; RedNote 88). Two raters evaluated overall quality (Global Quality Score, GQS 1–5), reliability (modified DISCERN, 0–5), source credibility (JAMA, 0–4), and educational usability (PEMAT-Understandability and- Actionability, 0–100%). Content completeness was coded across six domains (epidemiology, etiology, clinical manifestations, diagnosis, treatment, prevention) to create a Content Completeness Index (CCI, 0–12). Between - group comparisons used non - parametric tests; proportional-odds ordinal logistic regression examined independent predictors of higher GQS and mDISCERN.

Results

Median video length was 123 s (IQR 66–254), with the longest on Bilibili (298 s). Platform differences were observed for overall quality ($P < 0.001$), reliability ($P = 0.040$), JAMA score ($P < 0.001$), and actionability ($P = 0.021$): TikTok and Bilibili scored higher on GQS than RedNote; Bilibili had higher reliability; TikTok showed the highest JAMA score and higher actionability than RedNote. Understandability did not differ across platforms. Licensed uploaders (physicians, licensed non-physicians, institutions) outperformed non-licensed accounts (health communicators, patient-caregivers) on GQS ($P < 0.0001$), mDISCERN ($P < 0.05$), JAMA score ($P < 0.0001$), and actionability ($P < 0.01$), with no difference in understandability. Quality, reliability, and usability increased stepwise with higher CCI (all $P < 0.001$). In adjusted models, CCI was the strongest predictor of higher GQS (OR 2.46 per 1-point increase; 95% CI 2.05–2.96) and higher mDISCERN (OR 1.26; 95% CI 1.08–1.47). Patient–caregiver uploads had lower reliability than institutional accounts (OR 0.09; 95% CI 0.01–0.70). Engagement metrics (likes, collections, comments) correlated strongly with each other but weakly with quality or reliability.

Conclusions

AD short videos on Chinese platforms show mixed quality with limited source transparency. Content completeness—not platform, popularity, or length—best predicts higher quality and reliability. Encouraging structured coverage of core domains, improving source disclosure, and promoting collaborations between licensed professionals and skilled communicators could raise the baseline quality of online mental - health education.

1. Introduction

Adjustment disorder is one of the most prevalent mental health conditions worldwide, characterized as a mild and often indistinct maladaptive response to non - traumatic stressors with considerable variability in symptomatology and functional impairment[1,2]. In DSM – 5, it was reconceptualized as a stress response syndrome with a trauma - related subtype, though its diagnostic boundaries remain debated[3,4]. The disorder is highly common in clinical settings, with prevalence estimates of 11%–18% in primary care and 10% – 35% in consultation - liaison psychiatry, accounting for 5%–20% of patient encounters[5]. Clinical presentations are diverse, including anxiety, depression, mixed emotional disturbances, or behavioral dysregulation, often accompanied by physiological symptoms such as insomnia, appetite changes, and palpitations[6]. According to ICD – 11, its defining features center on two core symptom clusters: preoccupation with the stressor and failure to adapt[4]. The direct health burden is reflected in disability - adjusted life years (DALYs), though precise attribution is complicated by frequent comorbidity[7,8]. Given its high prevalence, functional impact, and potential to progress to persistent mental disorders, early identification and intervention should be a public health priority.

In the digital era, social media has transformed how the public accesses health information, with short-video platforms becoming particularly influential[9,10].China's short-video ecosystem is highly heterogeneous[11]. TikTok is driven by algorithm-curated, short-form content; Bilibili accommodates longer, community-supported explainers; and RedNote integrates lifestyle-oriented, user-generated narratives[12,13]. These platform-specific norms of length, style, and audience expectations may differentially shape the quality and completeness of mental-health information. Prior evaluations on Chinese platforms show variable information quality by disease area and uploader identity;however, adjustment-disorder (AD)-specific analyses are lacking[14,15]. We conduct a cross-platform assessment of Mandarin-language AD videos on Tiktok, Bilibili, and Rednote using established quality and reliability instruments—the Global Quality Scale (GQS), modified DISCERN, JAMA benchmark criteria, and the Patient Education Materials Assessment Tool for audiovisual materials (PEMAT-A/V)—alongside a Content Completeness Index (CCI) covering six domains: definition/diagnostic criteria, symptomatology, etiology/stressors, risk/timeline/prognosis, evidence-based management, and guidance on help-seeking. We examine: (1) How do quality, reliability, understandability/actionability, and completeness differ by platform and uploader? (2) Which factors independently predict

higher quality and reliability? (3) Do engagement signals align with objective quality metrics? This framework seeks to inform clinicians, creators, and platforms on pragmatic strategies to strengthen online mental-health education.

2. Methods

2.1 Study Design and Ethical Considerations

This investigation was conducted as a cross - sectional analysis of publicly accessible short video content concerning adjustment disorder across three prominent Chinese social media platforms[16,17]. Given that the research exclusively utilized data available in the public domain and did not involve human subjects or confidential user information, it was exempt from formal review by an institutional ethics committee.

2.2 Eligibility and sampling

On March 20, 2026, a systematic search was performed on TikTok (Douyin), Bilibili, and RedNote (Xiaohongshu) using the Chinese search term "□□□□" (adjustment disorder). Supplementary Table 1 provides a detailed description of these platforms, including their distinct operational dynamics and user interaction features. To minimize the influence of personalization algorithms and ensure the objectivity of our search results, we employed newly created accounts with all browsing history, cookies, and cache cleared prior to data collection. For RedNote, which requires user login to access its search functionality, a new account was established using an unactivated mobile number to eliminate any potential impact from pre-existing user data and maintain a standardized search environment[18].

To mirror typical user exposure, we retrieved the top 100 items from each platform using the default comprehensive ranking. This strategy reflects established behavior patterns showing that about 70% of people seeking health information review only the first two pages of results, and over 90% of clicks occur on the first page[19,20]. Eligibility required Mandarin-language content with adjustment disorder as the primary focus. We excluded materials judged to be irrelevant, primarily self-promotional, reuploads, formal online medical courses, or duplicates. After screening, the initial pool of 300 was reduced to 228 for in-depth analysis (TikTok: 78; Bilibili: 62; RedNote: 88) (Fig. 1).

2.3 Video characteristics and uploader types

For each video, we recorded the following general characteristics: source platform, duration in seconds, days since publication, and key engagement metrics (likes, comments, collections, shares).

In contrast to the physician-led, procedure-centred model that typifies many medical specialties, mental health care is inherently multi-professional, with psychiatrists working alongside clinical psychologists, licensed psychotherapists or counselors, psychiatric nurses, and social workers who not only deliver assessment, intervention, and follow-up but also play a prominent role in public education[21,22]. To reflect this ecology and its implications for information quality, we defined uploader categories with finer granularity: physicians (licensed medical doctors), licensed non-physician mental-health professionals (clinical psychologists, licensed counselors/psychotherapists, psychiatric nurses and social workers), health communicators (individuals or organizations producing science communication without clinical licensure), patient–caregivers (self-identified patients or family members sharing lived experience), and healthcare institutions (official hospital or clinic accounts). Classification followed a verifiability-first rule grounded in publicly available evidence—account biographies, platform verification badges, institutional affiliations, license numbers or professional titles, and institutional webpages; where multiple identities were claimed, the highest level of clinical licensure took precedence; self-described "therapists" or "counselors" without verifiable credentials were coded as health communicators; official hospital or clinic accounts were coded as institutions; and personal narratives without professional credentials were coded as patient-caregivers[23]. These procedures were prespecified in the coding manual, piloted during rater calibration, and applied consistently to limit misclassification; we did not adjudicate the legal validity of credentials beyond what could be verified in the public domain. Framed this way, the taxonomy captures the distinctive, collaborative nature of mental health services and allows our analyses to isolate how provenance—professional licensure and institutional affiliation, rather than platform or popularity—shapes the quality, reliability, and educational value of short-video content on adjustment disorder.

2.4 Quality, reliability, and education assessments

To minimise measurement bias and enhance reproducibility, two independent raters evaluated all included videos while blinded to each other's scores and to engagement metrics (likes, comments, collections, shares); during the pilot phase they were also masked to platform identity, although full blinding in the main coding was not feasible because platform-specific interface cues were visible. Before main coding, a three-round calibration on a separate pilot set aligned interpretation and rules: round one standardised the codebook and operational definitions ; round two involved independent scoring followed by reconciliation and refinement of decision rules; round three confirmed consensus, after which any remaining disagreements during main coding were resolved by discussion and the consensus scores were used for all analyses.

In the scoring content, we examined four complementary dimensions—overall quality, reliability, educational value, and the breadth of medical content—to capture distinct facets of video performance. Established instruments underpinned this assessment: overall quality was captured with the 5-point Global Quality Score (GQS)[24]; reliability was evaluated using the five-item modified DISCERN and the four-item JAMA benchmark checklist[25,26]; and educational value was quantified with PEMAT-Understandability and PEMAT-Actionability, reported as percentages from 0 to 100[27]. Content breadth was summarized with the Content Completeness Index (CCI; 0–12) derived from six domains specified in Section 2.3, separating what topics are

covered from how they are presented or who presents them. Results are organized as follows: platform- and uploader-level comparisons of characteristics, engagement, and quality/reliability/content metrics appear in Tables 1–2; variation in quality and reliability across CCI categories—Low (0–3), Medium (4–5), and High (6–12)—is shown in Table 3; and factors associated with higher GQS and mDISCERN scores are modeled in the ordinal logistic regressions in Tables 4–5.

Table 1
Platform-level comparison of short-video characteristics, engagement, quality/reliability, and content metrics across TikTok, Bilibili, and RedNote

Variables	Total (n = 228)	TikTok (n = 78)	BiliBili (n = 62)	RedNote (n = 88)	P-value
General information					
Length (s), M (Q_1, Q_3)	123.00 (66.25, 254.00)	107.50 (62.00,166.25)	298.50 (147.75,596.25)	98.00 (48.25,199.75)	< .001
Length (min), M (Q_1, Q_3)	2.39 (1.21,4.13)	2.16 (1.25,3.07)	4.97 (2.46,9.94)	1.64 (0.81,3.33)	< .001
Days since published (days), M (Q_1, Q_3)	767.00 (317.00,1077.25)	966.00 (642.25,1186.75)	539.00 (155.75,1306.00)	529.00 (252.50,882.00)	< .001
Likes, M (Q_1, Q_3)	23.00 (6.00, 193.00)	23.00 (9.00,92.00)	22.50 (2.75,531.00)	25.50 (4.00,416.75)	< .912
Collections, M (Q_1, Q_3)	9.00 (2.00, 98.50)	5.00 (1.75,34.25)	27.50 (2.00,396.75)	10.50 (2.00,98.50)	0.006
Comments, M (Q_1, Q_3)	2.00 (0.00,13.50)	2.50 (0.00,9.25)	2.00 (0.00,36.25)	1.00 (0.00,13.50)	0.319
Shares, M (Q_1, Q_3)					
Uploader classification, n(%)					< .001
Physician	52 (22.81)	35 (44.87)	10 (16.13)	7 (7.95)	
Licensed Non-Physician	56 (24.56)	20 (25.64)	20 (32.26)	16 (18.18)	
Health Communicator	98 (42.98)	20 (25.64)	28 (45.16)	50 (56.82)	
Patient-Caregiver	16 (7.02)	1 (1.28)	2 (3.23)	13 (14.77)	
Healthcare Institution	6 (2.63)	2 (2.56)	2 (3.23)	2 (2.27)	
Quality and reliability					
GQS, M (Q_1, Q_3)	3.00 (2.00, 4.00)	3.00 (3.00,4.00)	3.00 (3.00,4.00)	3.00 (2.00,3.00)	< .001
mDISCERN, M (Q_1, Q_3)	2.00 (2.00, 2.00)	2.00 (2.00,2.25)	2.00 (2.00,3.00)	2.00 (2.00,2.00)	0.040
JAMA, M (Q_1, Q_3)	1.00 (1.00, 2.00)	2.00 (1.00,2.00)	1.00 (1.00,2.00)	1.00 (1.00,2.00)	< .001
PEMAT-U, M (Q_1, Q_3)	67.00 (56.00, 82.00)	70.00 (55.00,83.00)	64.00 (56.00,80.00)	67.00 (56.00,80.00)	0.367
PEMAT-A, M (Q_1, Q_3)	67.00 (0.00, 91.75)	67.00 (33.00,100.00)	33.00 (33.00,67.00)	50.00 (0.00,67.00)	0.021
Content analysis					
Content Completeness Index, M (Q_1, Q_3)	4.00 (2.00, 6.00)	4.00 (3.00,6.00)	4.50 (3.00,6.00)	2.50 (2.00,4.00)	< .001
Management modality					
Psychotherapy, n(%)	115 (50.44)	37 (47.44)	32 (51.61)	46 (52.27)	0.806
Pharmacological Treatment, n(%)	65 (28.51)	30 (38.46)	23 (37.10)	12 (13.64)	0.545
Somatic Therapies, n(%)	28 (12.28)	8 (10.26)	10 (16.13)	10 (11.36)	< .001

M = median; Q1/Q3 = first/third quartile; n(%) = frequency (percentage). P-values from Kruskal–Wallis tests for continuous/ordinal variables and Pearson's χ^2 (or Fisher's exact where appropriate) for categorical variables. Content Completeness Index (CCI) is the sum of six domains (0–12).

Table 2
Uploader-level comparison of short-video characteristics, engagement, quality/reliability, and content metrics (Doctors vs Science Communicators vs Patients & Family)

Variables	Total (n = 228)	Physician (n = 52)	Licensed Non-Physician (n = 56)	Health Communicator (n = 98)	Patient-Caregiver (n = 16)	Healthcare institution (n = 6)	P-value
General information							
Length (s), M (Q_1, Q_3)	123.00 (66.25, 254.00)	103.00 (51.75,161.00)	195.00 (81.00,518.00)	116.00 (64.00,212.00)	199.00 (105.25,313.25)	106.00 (56.25,128.75)	0.005
Length (min), M (Q_1, Q_3)	2.39 (1.21,4.13)	1.84 (0.91,3.16)	3.00 (1.43,7.92)	2.47 (1.17,3.70)	3.32 (1.76,5.22)	1.77 (0.94,2.15)	0.023
Days since published (days), M (Q_1, Q_3)	767.00 (317.00,1077.25)	1019.00 (499.50,1231.00)	647.00 (178.00,1208.00)	647.00 (310.00,1026.50)	424.50 (113.50,981.00)	746.50 (183.50,891.75)	0.014
Likes, M (Q_1, Q_3)	23.00 (6.00, 193.00)	37.00 (12.00,119.50)	16.00 (4.00,65.00)	24.00 (5.00,222.00)	409.00 (17.75,3056.00)	4.50 (1.50,555.75)	0.010
Collections, M (Q_1, Q_3)	9.00 (2.00, 98.50)	9.00 (2.25,55.50)	8.00 (2.00,69.00)	9.00 (2.00,116.00)	197.00 (8.00,685.00)	0.00 (0.00,282.50)	0.044
Comments, M (Q_1, Q_3)	2.00 (0.00,13.50)	4.00 (0.25,15.25)	1.00 (0.00,7.00)	1.00 (0.00,11.00)	16.00 (1.00,207.75)	0.00 (0.00,307.75)	0.017
Shares, M (Q_1, Q_3)							
Quality and reliability							
GQS, M (Q_1, Q_3)	3.00 (2.00, 4.00)	3.00 (3.00,4.00)	3.00 (3.00,4.00)	3.00 (2.00,3.00)	2.00 (1.00,3.00)	2.50 (1.75,4.00)	0.001
mDISCERN, M (Q_1, Q_3)	2.00 (2.00, 2.00)	2.00 (2.00,3.00)	2.00 (2.00,3.00)	2.00 (2.00,2.00)	2.00 (1.00,2.00)	2.00 (2.00,3.00)	0.025
JAMA, M (Q_1, Q_3)	1.00 (1.00, 2.00)	2.00 (2.00,2.00)	2.00 (1.00,2.00)	1.00 (1.00,1.00)	1.00 (1.00,2.00)	1.00 (1.00,2.00)	< .001
PEMAT-U, M (Q_1, Q_3)	67.00 (56.00, 82.00)	71.00 (56.25,83.00)	71.00 (56.00,83.00)	67.00 (55.50,75.00)	67.00 (47.00,67.00)	73.50 (56.00,93.25)	0.090
PEMAT-A, M (Q_1, Q_3)	67.00 (0.00, 91.75)	67.00 (33.00,100.00)	67.00 (33.00,100.00)	33.00 (0.00,67.00)	33.00 (0.00,100.00)	67.00 (24.75,100.00)	0.249
Content analysis							
Content Completeness Index, M (Q_1, Q_3)	4.00 (2.00, 6.00)	4.00 (3.00,7.00)	4.00 (3.00,6.00)	4.00 (2.00,5.00)	2.00 (1.00,3.00)	4.00 (1.75,5.25)	0.003
Management modality							
Psychotherapy, n(%)	115 (50.44)	33 (63.46)	24 (42.86)	42 (42.86)	11 (68.75)	5 (83.33)	0.019
Pharmacological Treatment, n(%)	65 (28.51)	24 (46.15)	15 (26.79)	24 (24.49)	2 (12.50)	0 (0.00)	0.011
Somatic Therapies, n(%)	28 (12.28)	9 (17.31)	4 (7.14)	13 (13.27)	2 (12.50)	0 (0.00)	0.477

M = median; Q1/Q3 = first/third quartile; n(%) = frequency (percentage). P-values from Kruskal–Wallis tests for continuous/ordinal variables and Pearson's χ^2 (or Fisher's exact where any expected cell count < 5) for categorical variables. Doctors include cardiology specialists, non-cardiology physicians, and TCM physicians (primary analysis); Science Communicators are personal or institutional science communication accounts; Patients & Family are self-reported patient/family accounts. Content Completeness Index (CCI) is the sum of six domains (0–12) scored 0/1/2 as defined.

Table 3
Video quality and reliability by levels of the Content Completeness Index (CCI)

Variables	Total (n = 228)	Low CII (n = 99)	Medium CII (n = 68)	High CII (n = 61)	P-value
GQS, M (Q_1, Q_3)	3.00 (2.00, 4.00)	2.00 (2.00,3.00)	3.00 (3.00,4.00)	4.00 (3.00,4.00)	< .001
mDISCERN, M (Q_1, Q_3)	2.00 (2.00, 2.00)	2.00 (2.00,2.00)	2.00 (2.00,3.00)	2.00 (2.00,3.00)	< .001
JAMA, M (Q_1, Q_3)	1.00 (1.00, 2.00)	1.00 (1.00,2.00)	1.00 (1.00,2.00)	2.00 (1.00,2.00)	0.025
PEMAT-U, M (Q_1, Q_3)	67.00 (56.00, 82.00)	60.00 (50.00,71.00)	68.50 (60.00,83.00)	73.00 (60.00,82.50)	< .001
PEMAT-A, M (Q_1, Q_3)	67.00 (0.00, 91.75)	33.00 (0.00,67.00)	67.00 (33.00,67.00)	67.00 (49.50,100.00)	< .001

Values are median (IQR). P values from Kruskal–Wallis tests (H statistic). Content Completeness Index (CCI) is the sum of six domains (0–12); groups defined as Low (0–3), Medium (4–5), High (6–12).

Table 4
Ordinal logistic regression of factors associated with higher Global Quality Score (GQS) in short-video content

Variables	β	S.E	z	OR (95%CI)	P-value
Platform					
RedNote				1.00 (Reference)	
TikTok	-0.05	0.36	-0.13	0.96 (0.48 ~ 1.92)	0.899
Bilibili	0.25	0.38	0.66	1.29 (0.60 ~ 2.74)	0.517
Classification					
Healthcare institution				1.00 (Reference)	
Physician	0.40	0.85	0.47	1.49 (0.30 ~ 7.59)	0.626
Licensed Non-Physician	1.02	0.84	1.21	2.78 (0.56 ~ 13.78)	0.210
Health Communicator	0.27	0.82	0.33	1.31 (0.28 ~ 6.17)	0.732
Patient–Caregiver	-0.09	0.94	-0.10	0.91 (0.15 ~ 5.63)	0.920
Length(minutes)	-0.02	0.03	-0..47	0.99 (0.93 ~ 1.05)	0.646
Days Since Published(days)	0.00	0.00	1.20	1.00 (1.00 ~ 1.00)	0.231
Content Completeness Index	0.90	0.09	9.79	2.46 (2.05 ~ 2.96)	< .001

Proportional odds ordinal logistic regression. Outcome = GQS (1–5). Reference groups: platform = RedNote; classification = Healthcare institution. Video length was rescaled to minutes (OR per 1-minute increase). Days since published in days. Content Completeness Index (CCI) range 0–12. OR > 1 denotes increased odds of higher GQS.

Table 5
Ordinal logistic regression of factors associated with higher mDISCERN reliability score in short-video content

Variables	β	S.E	z	OR (95%CI)	P-value
Platform					
RedNote				1.00 (Reference)	
TikTok	-0.62	0.42	-1.49	0.54 (0.24 ~ 1.22)	0.137
Bilibili	0.07	0.42	0.16	1.07 (0.47 ~ 2.42)	0.870
Classification					
Healthcare institution				1.00 (Reference)	
Physician	-0.35	0.91	-0.38	0.71 (0.12 ~ 4.08)	0.701
Licensed Non-Physician	-0.77	0.91	-0.85	0.46 (0.08 ~ 2.67)	0.395
Health Communicator	-1.11	0.89	-1.25	0.33 (0.06 ~ 1.82)	0.211
Patient–Caregiver	-2.43	1.07	-2.28	0.09 (0.01 ~ 0.70)	0.023
Length(minutes)	0.04	0.03	1.31	1.04 (0.98 ~ 1.11)	0.189
Days Since Published(days)	0.00	0.00	0.59	1.00 (1.00 ~ 1.00)	0.557
Content Completeness Index	0.23	0.08	2.96	1.26 (1.08 ~ 1.47)	0.003

Proportional odds ordinal logistic regression. Outcome = mDISCERN (0–5). Reference groups: platform = RedNote; classification = Healthcare institution. Video length was rescaled to minutes (OR per 1-minute increase). Days since published in days. Content Completeness Index (CII) range 0–12. OR > 1 denotes increased odds of higher mDISCERN.

2.5 Statistical analysis

Continuous/ordinal variables were summarized as medians (IQR) and compared with Kruskal–Wallis tests. Categorical variables were compared with χ^2 or Fisher's exact tests. We used proportional - odds ordinal logistic regression to identify independent predictors of higher GQS and higher mDISCERN, modeling platform, uploader classification, video length (minutes), days since publication (days), and CCI (0–12). Reference categories matched Table 4–5. Two - sided P<.05 was considered statistically significant. Analyses were conducted using standard statistical software.

3. Results

3.1 Sample characteristics

Screening yielded 228 unique Mandarin videos on adjustment disorder (TikTok 78; Bilibili 62; RedNote 88) from 300 top-ranked search results (Fig. 1). Median length was 123 s (IQR 66–254) and was longest on Bilibili (298.5 s) compared with TikTok (107.5 s) and RedNote (98.0 s) ($P < 0.001$; Table 1). Videos on TikTok were older (days since publication median 966) than those on RedNote (529) and Bilibili (539) ($P < 0.001$). Platform-level engagement differed for collections ($P = 0.006$) but not for likes ($P = 0.912$) or comments ($P = 0.319$); shares were inconsistently available across platforms and were not used in the primary comparisons. Uploader mix varied by platform ($P < 0.001$): health communicators predominated overall (42.98%), followed by licensed non - physicians (24.56%), physicians (22.81%), patient–caregivers (7.02%), and healthcare institutions (2.63%) (Table 1; Fig. 2).

3.2 Platform - level comparisons of quality, reliability, and education

Overall quality (GQS) differed across platforms ($P < 0.001$). Pairwise comparisons (Fig. 3A) showed TikTok and Bilibili scoring higher than RedNote (both $P < 0.01$), with no significant difference between TikTok and Bilibili. Medians: TikTok 3.0 (3.0–4.0); Bilibili 3.0 (3.0–4.0); RedNote 3.0 (2.0–3.0). Reliability (mDISCERN) differed modestly ($P = 0.040$), driven by higher scores on Bilibili versus TikTok and RedNote (both $P < 0.01$), while TikTok and RedNote did not differ (Fig. 3B). Source credibility (JAMA) varied markedly ($P < 0.001$), with TikTok outperforming RedNote ($P < 0.0001$) and Bilibili ($P < 0.01$), whereas RedNote and Bilibili were similar (Fig. 3C). Understandability (PEMAT - U) was comparable across platforms ($P = 0.367$; Fig. 3E). Actionability (PEMAT - A) differed ($P = 0.021$), with TikTok exceeding RedNote ($P < 0.05$) and other platform pairs not significant (Fig. 3D). Content completeness (CCI) was higher on Bilibili and TikTok than RedNote (medians 4.5 and 4.0 vs 2.5; $P < 0.001$; Table 1). Domain-level coverage by platform (Fig. 5A) showed significant differences in etiology ($P < 0.0001$), clinical manifestations ($P < 0.05$), diagnosis ($P < 0.05$), and prevention ($P < 0.001$), whereas epidemiology and treatment did not differ significantly.

3.3 Uploader - level comparisons and management modalities

Quality and reliability differed by uploader (Table 2; Figs. 4A–E). Patient–caregiver videos had the lowest GQS (median 2.0; overall $P = 0.001$) and lower reliability (mDISCERN overall $P = 0.025$), whereas physicians and licensed non-physicians performed better on JAMA (both medians 2.0 vs 1.0 for

health communicators; P < 0.001). When uploaders were grouped as licensed (physicians, licensed non - physicians, institutions) versus non - licensed (health communicators, patient–caregivers), licensed videos scored higher on GQS (P < 0.0001), mDISCERN (P < 0.05), JAMA (P < 0.0001), and PEMAT - Actionability (P < 0.01), with no difference in PEMAT - Understandability (Fig. 6). Domain - level coverage by uploader (Fig. 5C) differed for etiology and diagnosis (both P < 0.01), favoring professional uploaders; other domains were comparable. Management modalities by platform showed broadly similar psychotherapy and medication mentions (P = 0.806 and P = 0.545, respectively) and significant variation in somatic therapies (P < 0.001), more frequent on Bilibili (Fig. 5B; Table 1). Descriptively, psychotherapy appeared in about half of the videos on each platform (Bilibili 51.6%, RedNote 52.3%, TikTok 47.4%); pharmacological treatment appeared in 37.1% on Bilibili, 38.5% on TikTok, and 13.6% on RedNote; somatic therapies in 16.1%, 10.3%, and 11.4%, respectively. By uploader type, psychotherapy and pharmacological treatment mentions varied significantly (P = 0.019 and P = 0.011, respectively; Table 2; Fig. 5D). Physicians were most likely to discuss psychotherapy (63.5%) and medication (46.2%), health communicators less so (42.9% and 24.5%), and patient–caregivers least likely to discuss medication (12.5%). Somatic-therapy mentions did not differ by uploader (P = 0.477).

3.4 Content completeness and monotonic gradients

Quality, reliability, and educational usability increased stepwise with higher CCI (Table 3). Median GQS rose from 2.0 (low CCI) to 3.0 (medium) and 4.0 (high) (P<.001). mDISCERN and JAMA were higher in medium/high CCI groups than low (both P<.001 and P = 0.025). Understandability (PEMAT - U) and actionability (PEMAT - A) also increased with completeness (both P<.001).

Correlation analysis (Fig. 7) showed strong collinearity among engagement metrics (likes-collections-comments ρ = 0.69–0.94) but weak or negligible correlations between engagement and quality/reliability/education scores (|ρ|≤0.15). Video length correlated modestly with GQS (ρ = 0.30), mDISCERN (ρ = 0.29), JAMA (ρ = 0.36), and PEMAT - U (ρ = 0.42).

3.5 Independent predictors of higher quality and reliability

In proportional - odds models (Tables 4–5), the Content Completeness Index (CCI) was the strongest independent predictor of higher GQS (OR 2.46 per 1 - point increase; 95% CI 2.05–2.96; P < 0.001) and higher mDISCERN (OR 1.26; 95% CI 1.08–1.47; P = 0.003). Platform, video length, and recency were not significant independent predictors after adjustment. Compared with institutional accounts, patient–caregiver uploaders had lower reliability (OR 0.09; 95% CI 0.01–0.70; P = 0.023). Other uploader categories did not differ significantly from institutions in adjusted models.

4. Discussion

4.1 Platform characteristics and content dissemination

Across platforms, adjustment-disorder videos showed moderate overall quality and generally limited provenance. In unadjusted comparisons, Bilibili hosted longer videos with higher mDISCERN (reliability) and greater content completeness (CCI), consistent with its community-supported, long-form orientation[25,26]. TikTok ranked highest on JAMA source credibility and had more actionable content (PEMAT-A) than RedNote, whereas RedNote's lifestyle emphasis coincided with shorter and less complete videos[13,28,29]. However, in multivariable models controlling for uploader type, video length, presence of credentials/references, and CCI, platform was not an independent predictor of GQS or mDISCERN. Taken together, these findings suggest that structured completeness and clear provenance cues—rather than the platform per se—drive higher quality and reliability. Platforms and creators could therefore prioritize completeness scaffolds such as checklists or chaptered sections covering key AD domains and ensure visible source cues including credentials, references, and update dates to improve quality without sacrificing usability[15].Linking our cross-platform patterns to known audience profiles and affordances may clarify mechanisms. TikTok's youth-skewed, identity-centric feed tends to reward concise narratives with salient provenance cues and explicit next steps, which aligns with its higher JAMA and actionability despite shorter durations[30,31]. Bilibili's community-moderated, explainer ecosystem attracts viewers willing to engage with longer formats, consistent with its higher completeness and mDISCERN[32]. RedNote's lifestyle and social-commerce orientation privileges personal storytelling and productized advice, which may contribute to shorter, less comprehensive coverage. These user portraits are descriptive rather than measured here; however, they provide a plausible mapping from platform environment to creator choices (format, sourcing, calls-to-action) and subsequently to quality and completeness outcomes.

4.2 The role of uploader identity in information quality

Uploader identity shaped credibility and completeness. Licensed professionals (physicians, licensed non - physicians, institutions) produced videos with higher GQS, mDISCERN, and JAMA and more actionable guidance than non - licensed uploaders; yet understandability did not differ, suggesting that clarity can be achieved by both groups. Patient - caregiver narratives drew engagement but had the lowest quality and reliability - valuable for peer support but insufficient for comprehensive education. Co - production models pairing clinicians with skilled communicators could combine credibility and engagement to deliver clearer, better - sourced content without sacrificing reach.

4.3 Quality disparities among physician uploaders

Within professional categories, physicians and licensed non-physicians performed comparably on GQS and mDISCERN, while institutional accounts showed relatively higher medians for understandability and actionability in figures, but had small numbers; institutional GQS medians were not higher than those of physicians. These patterns suggest professional training and organizational processes can raise baseline credibility, yet structured coverage is still necessary to translate expertise into higher perceived quality.

4.4 Discrepancies between video content and clinical guidelines

Across platforms, management segments emphasized psychotherapy, aligning with first-line care for adjustment disorder, while mentions of pharmacotherapy were less frequent overall and varied by uploader; physicians were most likely to include medications, whereas patient–caregiver uploads seldom did. Platform-level differences in medication mentions were not statistically significant in this sample, though descriptively RedNote featured fewer such mentions. Somatic therapies were rarely discussed, consistent with their limited role in adjustment disorder[33]. Beyond frequency, several guideline-salient elements were inconsistently covered: videos often provided limited clarification of diagnostic boundaries and differential diagnosis and variably addressed risk stratification ; indications for initiating pharmacotherapy, expected duration and discontinuation, and monitoring for adverse effects and drug–drug interactions were seldom specified; adherence strategies and follow-up intervals were inconsistently described; and few videos explicitly delineated when self-help or psychotherapy alone may be insufficient. These gaps risk misclassification and missed high-risk presentations, foster uncertainty about when and how to use medications for targeted symptoms, and may inadvertently encourage reliance on non–evidence-based remedies. To support guideline-concordant care, future videos should clarify diagnostic boundaries, time course, functional impairment criteria, and common differentials, including red flags and when to seek urgent care; incorporate structured risk stratification; emphasize evidence-based psychotherapies with clear indications, formats, and expected timelines for benefit; specify when pharmacotherapy may be appropriate for targeted symptoms or comorbidity, including initiation thresholds, expected onset of action, recommended duration, tapering, and situations to avoid or use caution; outline safety monitoring, follow-up cadence, and criteria for stepping up care or referral; provide concrete, actionable steps and link to validated self-help resources; explicitly state scenarios where psychotherapy and self-help alone are insufficient and indicate pathways for higher-intensity care; and encourage adherence through expectation-setting, normalizing help-seeking, and practical engagement tips while recommending evidence-based resources. Public health agencies and professional societies could develop concise, standardized explainers on adjustment disorder diagnosis, risk assessment, and stepped care, and collaborate with platforms to verify and label such content to improve discoverability and trust[34].

4.5 The decoupling of engagement metrics from content quality

Engagement indicators were highly intercorrelated yet largely uninformative about quality, reliability, or usability. Length correlated modestly with quality and usability but did not predict higher scores after adjusting for completeness. Together with the strong, monotonic gradients across CCI strata and the regression findings, these data indicate that structured, comprehensive coverage—not popularity or duration—drives educational value.Commercial incentives likely intersect with these engagement dynamics. In attention markets where conversions to private consultations or paid courses are monetized, creators may lean on anxiety-framed hooks (urgency, expansive symptom lists, “act-now” framing) that drive likes and saves but add little to reliability or completeness. The weak correlations between engagement and quality in our data are consistent with this decoupling: affective salience can generate reach even when sourcing is sparse or content is narrow.

4.6 Practical implications and recommendations

At the creation stage, adopt a consistent structured outline that covers epidemiology, etiology, clinical features, diagnosis, treatment, and prevention; disclose authorship, credentials, and last updated dates; cite credible sources and provide clear action steps and care pathways; and in short formats preserve completeness through tight structuring using concise chapter markers or on-screen checklists. Within professional channels, maintain rigorous source transparency to sustain stronger JAMA and mDISCERN performance and pair licensed clinicians with skilled communicators to maximize clarity and engagement without narrowing scope. At the platform level, embed low friction safeguards into workflows by adding dedicated fields for credentials and citations when diagnostic claims are made, standardizing last updated and evidence source tags, prompting inclusion of helplines and referral options, enforcing clear labeling of sponsored material, and integrating lightweight completeness checklists into upload pipelines and recommendation logic to ensure coverage of definitions, diagnostic thresholds, management options, and indications for seeking care. Collectively, these measures reconcile reach with reliability and provide a practical path to higher quality public facing mental health content.

4.7 Limitations

This study has several limitations. First, regarding measurement tools, the GQS, mDISCERN, and JAMA scores exhibited limited variability, with comparable medians across different platforms. This finding could indicate a true underlying similarity in content quality—implying that most short-form health videos cluster around a "moderate" level—but it may also reflect insufficient sensitivity of the rating scales. Although validated, the 5-point Likert scales might not capture fine-grained differences, potentially producing ceiling or floor effects that obscure more subtle cross-platform variations. Moreover, despite using three independent evaluators to standardize the assessment, these instruments are inherently subjective, introducing systematic bias that cannot be fully eliminated. Additionally, the classification of psychology-related videos was often ambiguous, making it difficult to distinguish between educational, entertainment, and therapeutic content.

Second, regarding study design and sampling, the cross-sectional approach allows for correlational inferences only, not causal conclusions. The strategy of selecting the top 100 videos, while reflecting common user browsing habits, may introduce ranking bias and is vulnerable to survivorship bias because deleted or lower-ranked content is excluded. Likewise, our chosen keyword set might have omitted semantically similar videos. Finally, the dynamic nature of digital platforms poses a challenge; algorithms are updated rapidly and continuously, so results obtained at a single time point may have limited generalizability over time.

5. Conclusions

AD short videos on Chinese platforms show mixed quality and limited source transparency. Structured content completeness—not platform, popularity, or length—best predicts higher quality and reliability. Promoting completeness checklists, improving source disclosure, and fostering collaborations between licensed professionals and skilled communicators may raise the baseline quality of online mental-health education.

Declarations

Acknowledgements

We extend our gratitude to the content creators whose videos on the three platforms were included in this analysis.

Funding

National Innovation and Entrepreneurship Training Program for Chinese College Students(202510344079X)

2024 National TCM Graduate Education Research Course Topics(YJS-YB-2024-92)

Author contributions

Y.J. conceived the study, developed the methodology, performed formal analysis and investigation, and drafted the main manuscript text. C.H. conducted investigation, validation, and data curation, co-drafted the manuscript, and prepared Figures 1–5. X.Y. contributed to methodology and visualization and co-drafted the manuscript. J.J. performed investigation and contributed to review and editing. H.L.Y. (Hongli Yang) and H.Y. (Han Yan) reviewed and edited the manuscript. L.C. prepared Figures 6–7 and assisted with visualization. X.C. contributed to conceptualization, supervision, project administration, and funding acquisition. L.L.C. (Liangliang Chen) provided overall supervision, critically revised the manuscript, and served as the corresponding author. All authors reviewed and approved the final manuscript.

Data availability

The datasets supporting the findings of this study, including the list of analyzed videos and their corresponding quality assessment scores, are available from the corresponding author upon reasonable request for the purposes of academic verification or secondary analysis.

Ethical approval

Ethical approval was not required for this study, as the data were sourced exclusively from publicly available videos on the TikTok, Bilibili, and RedNote platforms. The research involved no direct interaction with human subjects, and all analyzed content was in the public domain. The study did not collect any personally sensitive information, clinical data, or human specimens. All data collection and analysis procedures were conducted in full compliance with the platforms' terms of service, and all data were anonymized to protect privacy.

Consent for publication

All authors agreed to publish in the journal.

Competing interests

All authors explicitly stated that there was no conflict of interest.

References

1. Levin Y, et al. Network structure of ICD-11 adjustment disorder: a cross-cultural comparison of three African countries. Br J Psychiatry. 2021;219:557–64. 10.1192/bjp.2021.46.
2. Fernández-Buendía S, et al. Technology-supported treatments for adjustment disorder: A systematic review and preliminary meta-analysis. J Affect Disord. 2024;347:29–38. 10.1016/j.jad.2023.11.059.
3. Friedman MJ, et al. Classification of trauma and stressor-related disorders in DSM-5. Depress Anxiety. 2011;28:737–49. 10.1002/da.20845.
4. Strain JJ. The psychobiology of stress, depression, adjustment disorders and resilience. World J Biol Psychiatry. 2018;19:S14–20. 10.1080/15622975.2018.1459049.
5. Casey P. Adjustment disorder: epidemiology, diagnosis and treatment. CNS Drugs. 2009;23:927–38. 10.2165/11311000-000000000-00000.
6. Lavandier A, et al. Definition and Conceptual Interrogation of Complex Post-Traumatic Stress Disorder: A Scoping Review. Trauma Violence Abuse. 2025;15248380251397415. 10.1177/15248380251397415.
7. Ohi K, et al. Is adjustment disorder genetically correlated with depression, anxiety, or risk-tolerant personality trait? J Affect Disord. 2023;340:197–203. 10.1016/j.jad.2023.08.019.
8. Maercker A, Brewin CR. Controversies in trauma- and stress-related disorders. Br J Psychiatry. 2025;226:7–9. 10.1192/bjp.2024.147.
9. Zhou D, Lin LM. Short-video engagement, health-related behaviors, and mental well-being: Evidence from a longitudinal study of Chinese youth. Acta Psychol (Amst). 2026;265:106624. 10.1016/j.actpsy.2026.106624.

10. Liu J, et al. How do health content creators perform well? An integration research of short video and livestream behaviors. Front Public Health. 2024;12:1446247. 10.3389/fpubh.2024.1446247.
11. Zhou S, Cao Y. Multi-stakeholder game dynamics in copyright governance of China's short video: A stakeholder theory perspective. Telecomm Policy. 2026;50:103202–103202. 10.1016/j.Telpol.2026.103202.
12. Chun W. Identity Construction in Intercultural Communication on Social Media: A Case Study of Two Influencers on REDnote. *SHS Web of Conferences* 220, 02016–02016. 10.1051/shsconf/202522002016 (2025).
13. Hao P, Liu G, Lian S, Huang J, Zhao L. Evaluating the quality of TikTok videos on coronary artery disease using various scales to examine correlations with video characteristics and high-quality content. Sci Rep. 2025;15:9189. 10.1038/s41598-025-93986-3.
14. Ding R, et al. Health information in short videos about metabolic dysfunction-associated steatotic liver disease: Analysing quality and reliability. Liver international: official J Int Association Study Liver. 2024;44:1373–82. 10.1111/liv.15871.
15. Guo F, Ding G, Zhang Y, Liu X. Quality Assessment of Radiotherapy Health Information on Short-Form Video Platforms of TikTok and Bilibili: Cross-Sectional Study. JMIR cancer. 2025;11:e73455. 10.2196/73455.
16. Zheng X, Li Q, Jin L, Shi K, Deng M. Quality, reliability, and dissemination of lung cancer information on short-video platforms in China: a cross-platform content analysis of TikTok, Kwai, and Rednote. Front Public Health. 2025;13:1683561–1683561. 10.3389/fpubh.2025.1683561.
17. Wang Q, et al. Video in Chinese short video sharing platforms as a source of information on sleep disorders: A cross-sectional content analysis study. Digit health. 2026;12:20552076261415944. 10.1177/20552076261415944.
18. Gao S, et al. Does popularity reflect quality? A study on the reliability and quality of CAR-T videos on Chinese short-video platforms. Digit health. 2026;12:20552076261429670. 10.1177/20552076261429670.
19. M M-MJ. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychology behavior: impact Internet multimedia virtual Real Behav Soc. 2004;7:497–510. 10.1089/cpb.2004.7.497.
20. Hamilton KA, Qi L. Search Fluency Mistaken for Understanding: Ease of Information Retrieval from the Internet Inflates Internal Knowledge Confidence. Social Media + Soc. 2023;9. 10.1177/20563051231195547.
21. Hatchman O, Hagan SB, Foster K, Pemo K, Alexander L. From Representation to Integration: Lived Experience in Mental Health Teams: A Qualitative Descriptive Study. Int J Ment Health Nurs. 2026;35:e70230. 10.1111/inm.70230.
22. Christien M, Hans K, L, M. C., Jeannette P. Frying eggs or making a treatment plan? Frictions between different modes of caring in a community mental health team. Sociol Health Illn. 2021;43:1581–97. 10.1111/1467-9566.13346.
23. Tu J, Zhang C, Zhang H, Liang L, He J. Evaluating the reliability and quality of knee osteoarthritis educational content on TikTok and Bilibili: A cross-sectional content analysis. Digit health. 2025;11:20552076251366390. 10.1177/20552076251366390.
24. Kara M, Ozduran E, Kara MM, Özbek İ C, Hancı V. Evaluating the readability, quality, and reliability of responses generated by ChatGPT, Gemini, and Perplexity on the most commonly asked questions about Ankylosing spondylitis. PLoS ONE. 2025;20:e0326351. 10.1371/journal.pone.0326351.
25. He W, et al. Quality of cerebral palsy videos on Chinese social media platforms. Sci Rep. 2025;15:13323. 10.1038/s41598-024-84845-8.
26. Zhu W, et al. Information quality of videos related to esophageal cancer on tiktok, kwai, and bilibili: a cross-sectional study. BMC Public Health. 2025;25:2245–2245. 10.1186/s12889-025-23475-9.
27. Shoemaker SJ, Wolf MS, Brach C, Development of the Patient Education Materials Assessment Tool (PEMAT). A new measure of understandability and actionability for print and audiovisual patient information. Patient Educ Couns. 2014;96:395–403. 10.1016/j.pec.2014.05.027.
28. Zhou Y, 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. 2025;12:1613526–1613526. 10.3389/fmed.2025.1613526.
29. Wang H, Zhang H, Cao J, Zhang F, Xiong W. Quality and content evaluation of thyroid eye disease treatment information on TikTok and Bilibili. Sci Rep. 2025;15:25134. 10.1038/s41598-025-11147-y.
30. Craig MJA, et al. It Listens to Me So I Feel Well and Connected: Investigating the Influence of TikTok users' Perceived Algorithm Responsiveness and (In)sensitivity on Well-Being Via Self-Determination. Social Media + Soc. 2026;12. 10.1177/20563051261417301.
31. Gitomer A, Barthelemy JA, Welles BF. Youth activism and (de)personalized remix on TikTok. Social Mov Stud. 2026;25:215–33. 10.1080/14742837.2024.2415672.
32. Haoliang W, et al. The Current State of Vascular Surgery Presence in Bilibili Video Platform of China. Front Surg. 2022;9:874113–874113. 10.3389/fsurg.2022.874113.
33. Kelber MS, et al. Systematic review and meta-analysis of predictors of adjustment disorders in adults. J Affect Disord. 2022;304:43–58. 10.1016/j.jad.2022.02.038.
34. Poremski D, Hariram J, Wong WK, Eu PW, Lee C. The longitudinal dispositions of people diagnosed with adjustment or severe stress disorders. BMC Psychiatry. 2024;24:457. 10.1186/s12888-024-05904-y.

Figures

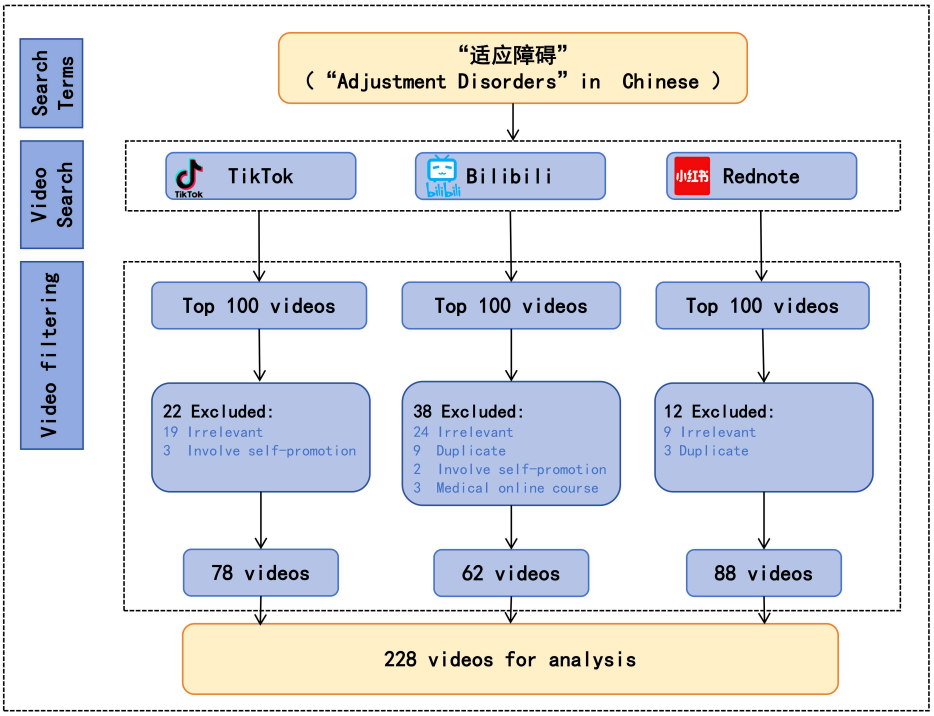


Figure 1

Flowchart of the video identification and selection process.

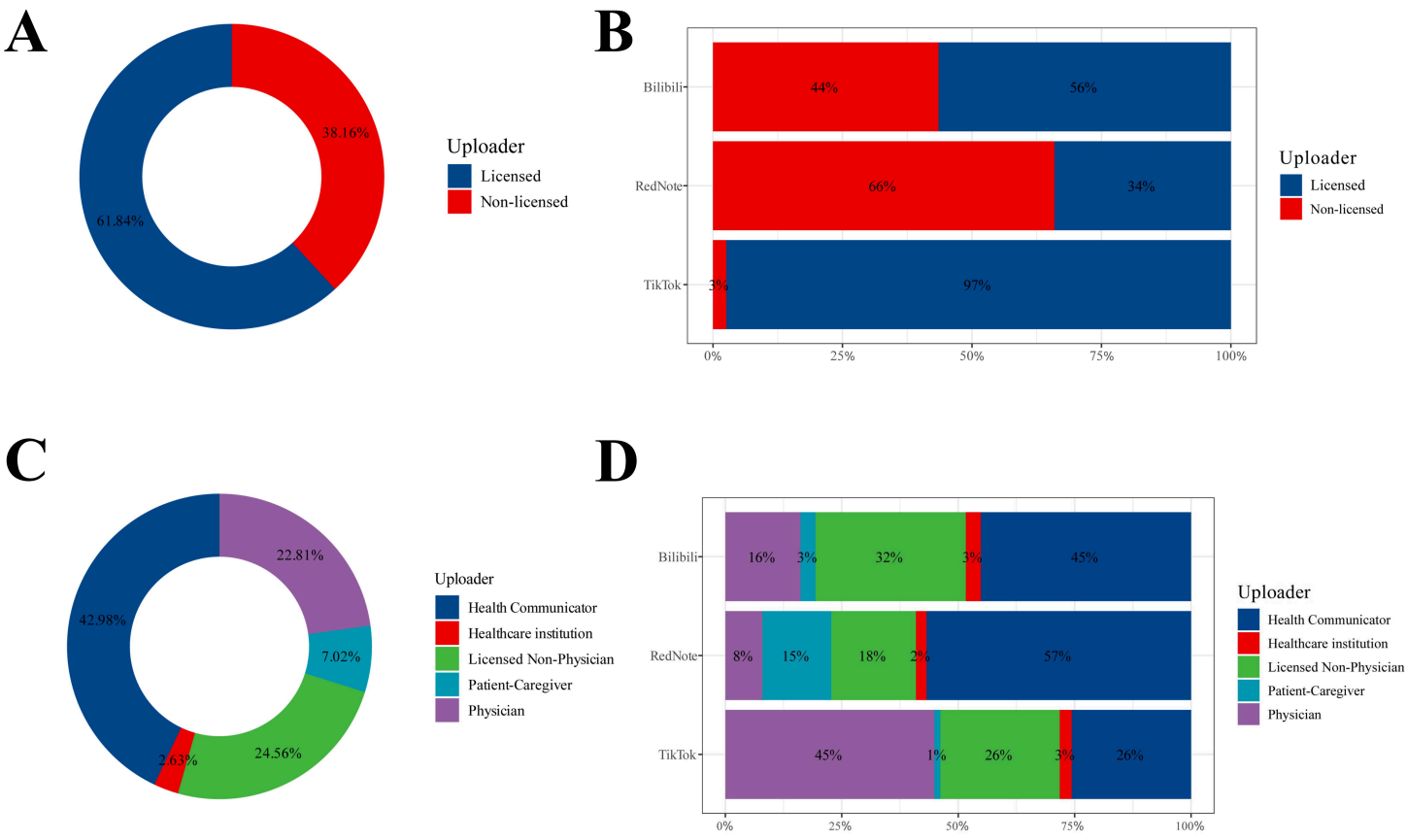


Figure 2

Distribution of video uploader identities. (A) Overall distribution of uploader types. (B) Distribution of uploader types across platforms. (C) Overall distribution of uploader types including physician subgroups. (D) Distribution of uploader types including physician subgroups across platforms.

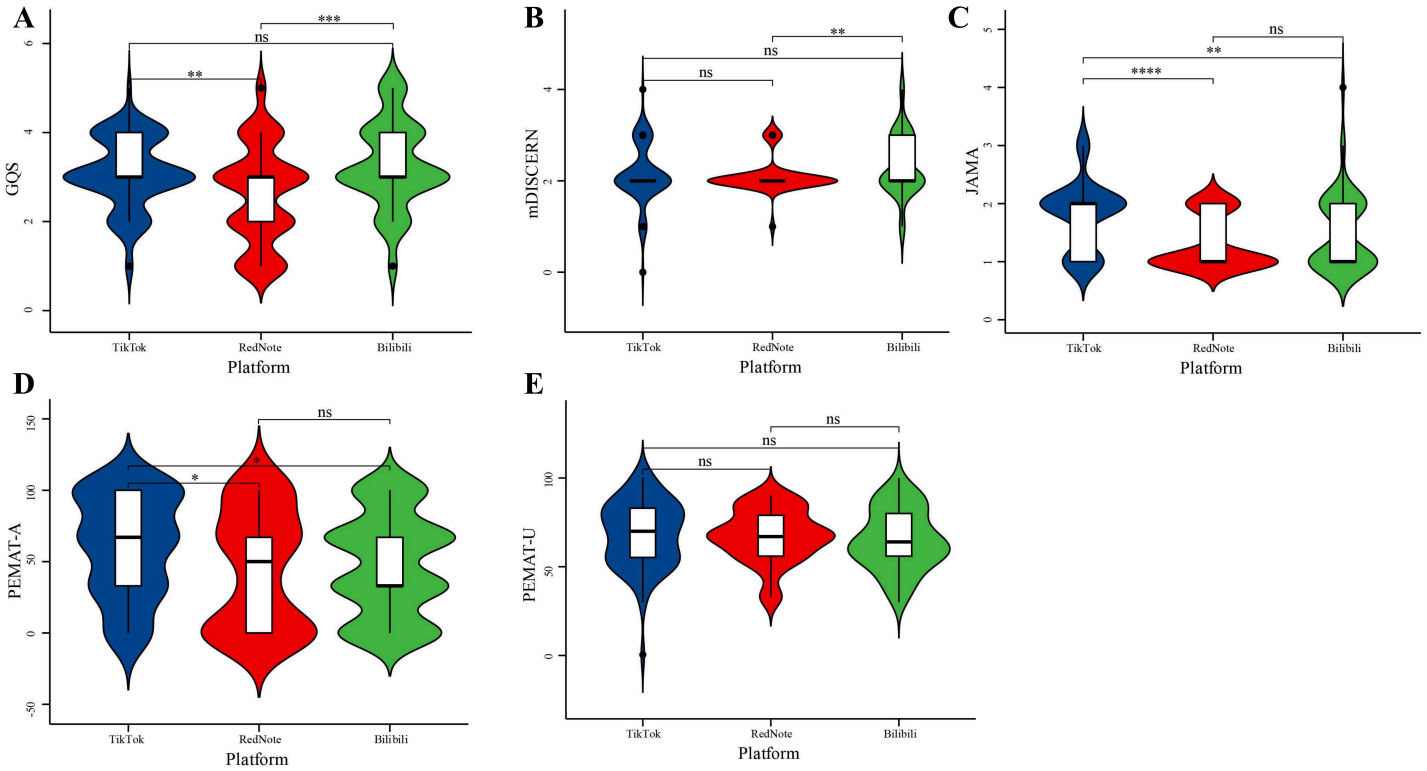


Figure 3

Comparison of video quality, reliability, and educational value scores across platforms.(A) GQS, (B) mDISCERN, (C) JAMA, (D) PEMAT-A, (E) PEMAT-U. Asterisks indicate statistically significant differences between groups (*P<.05, **P<.01). ns indicates no significant difference.

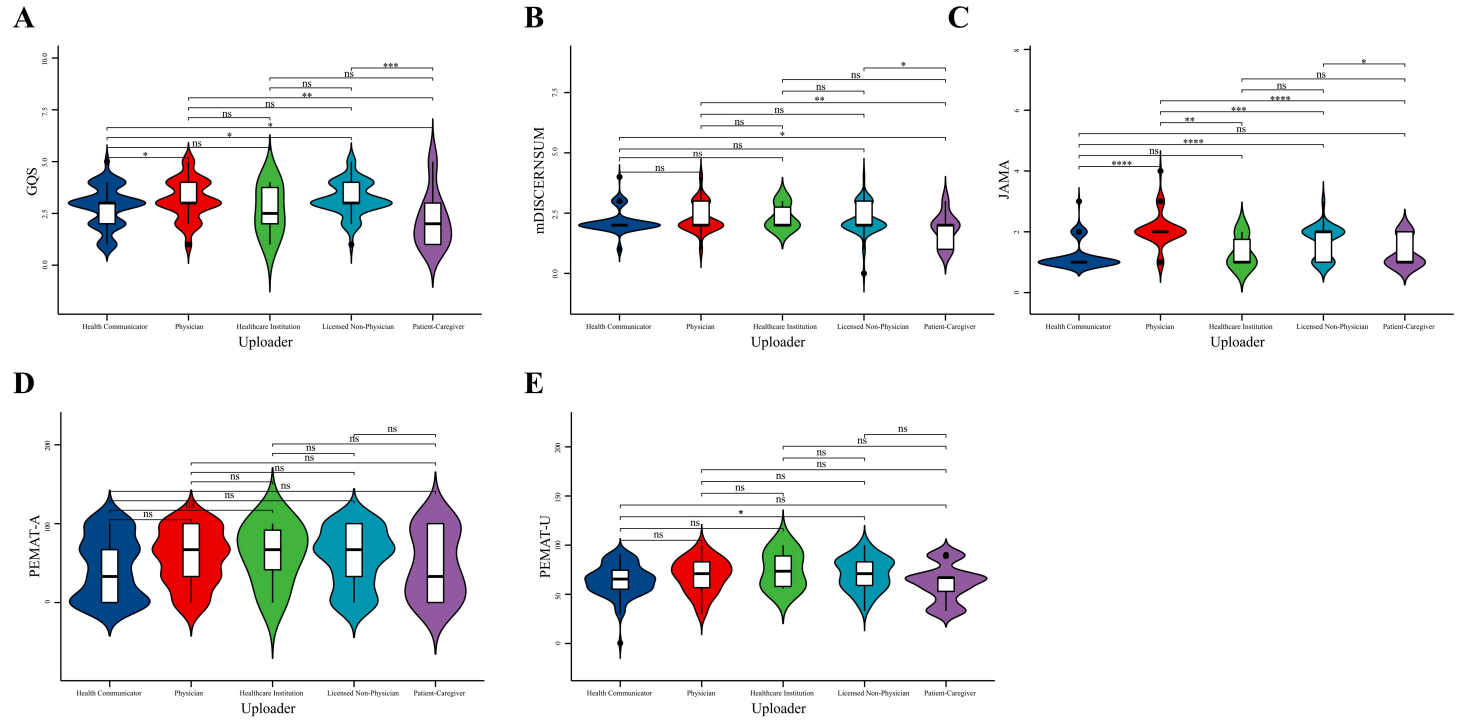


Figure 4

Comparison of video quality, reliability, and educational value scores by uploader type. (A) GQS, (B) mDISCERN, (C) JAMA, (D) PEMAT-A, (E) PEMAT-U. Asterisks indicate statistically significant differences between groups (*P<.05, **P<.01, ***P<.001). ns indicates no significant difference.

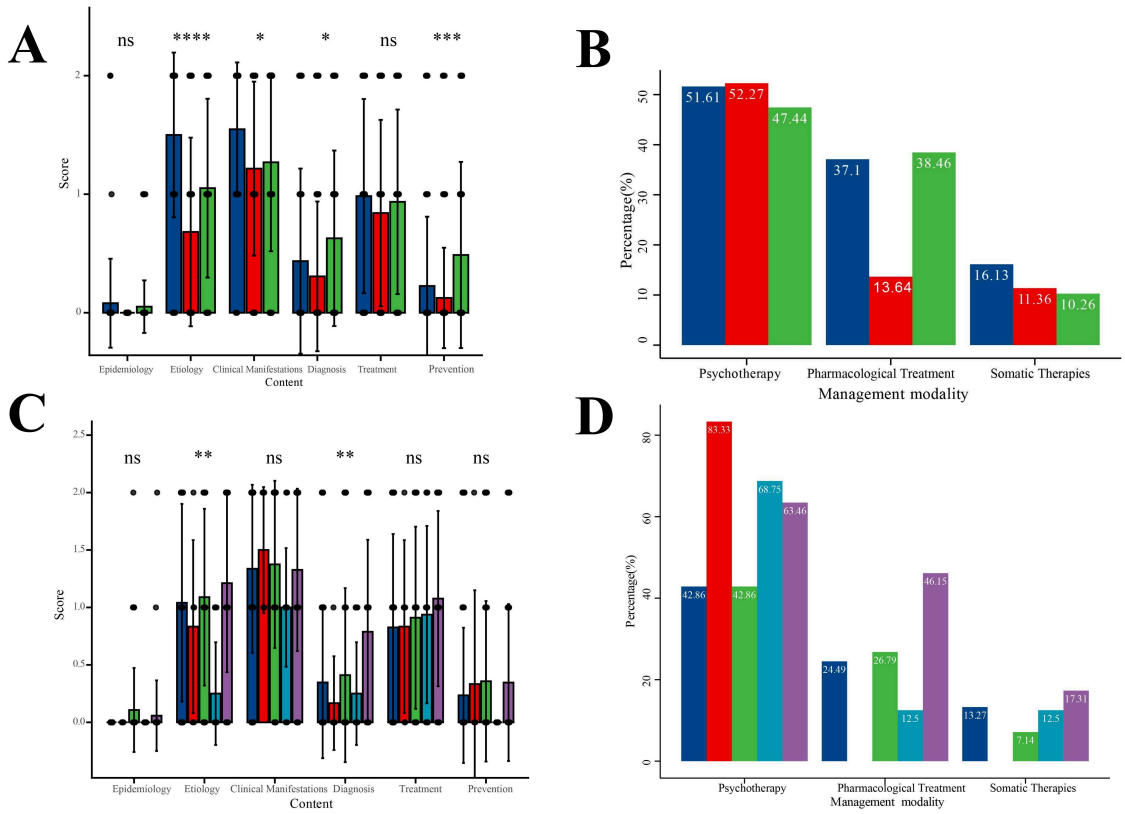


Figure 5

Comparison of video quality, reliability, and educational value scores among physician subgroups.(A) GQS, (B) mDISCERN, (C) JAMA, (D) PEMAT-A, (E) PEMAT-U. Asterisks indicate statistically significant differences between groups (**P<.01, ***P<.001). ns indicates no significant difference.

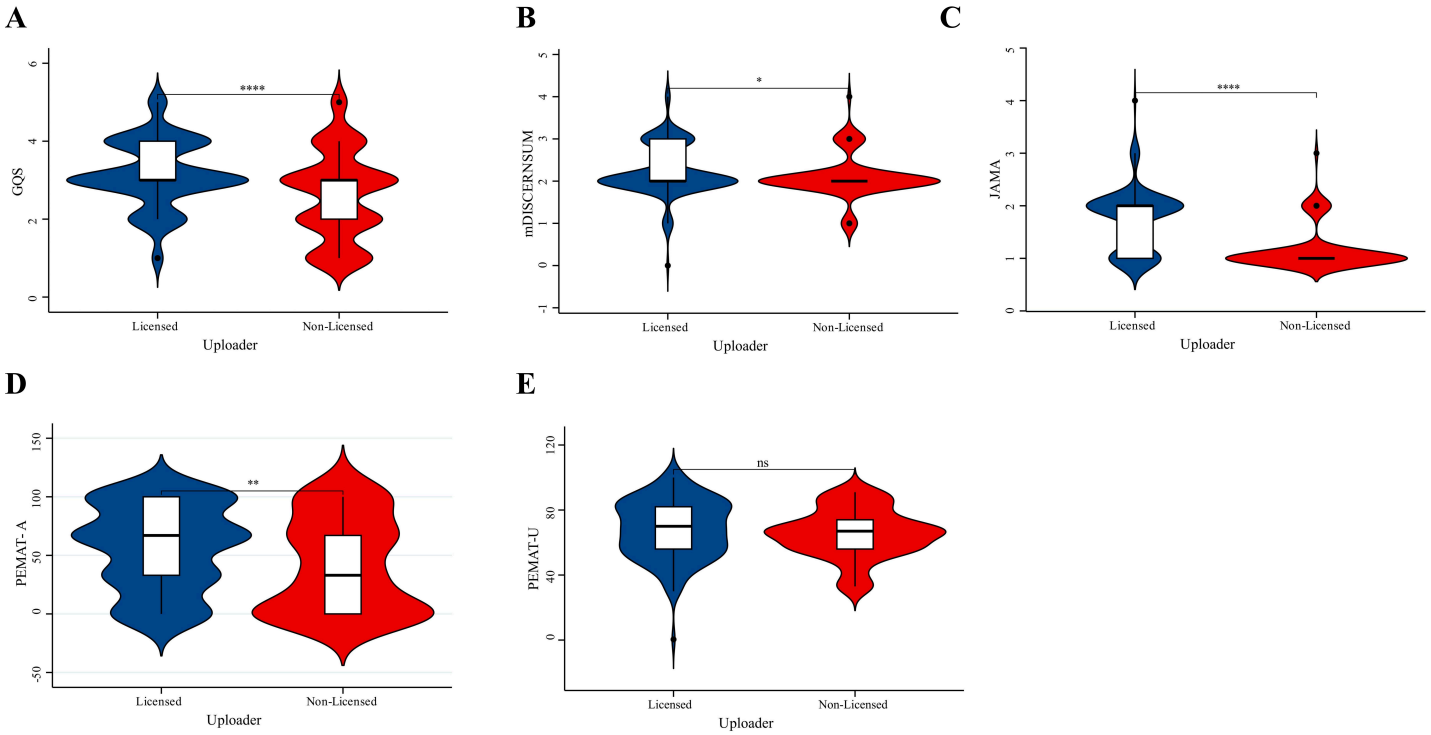


Figure 6

Analysis of video content.(A) Content domain coverage scores by platform. (B) Mention rates of management modalities by platform. (C) Content domain coverage scores by uploader type. (D) Mention rates of management modalities by uploader type.

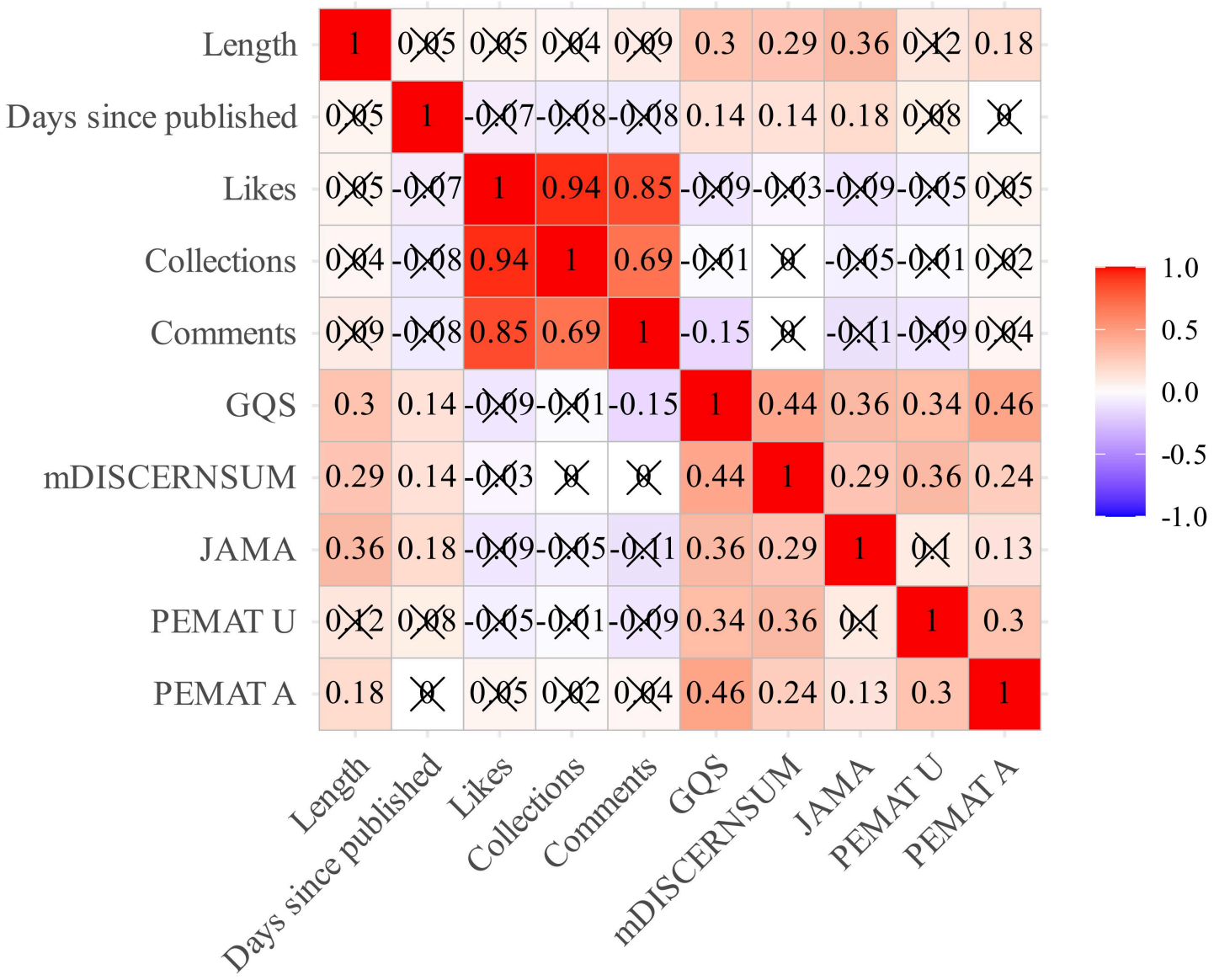


Figure 7

Spearman correlation matrix of video characteristics, user engagement metrics, and quality scores.