

ORIGINAL RESEARCH ARTICLE

Quality and Reliability of Chinese-Language Lumbar Disc Herniation Videos on TikTok and Bilibili: A Cross-Sectional Multiplatform Study

Chen Wang,^{#,†} Chao Yan,^{#,†,¶} Yibo Ma,[§] Guoyi Wang,^{l0} Yuyu Zhang,[⊥] Yayan Wang,[⊥] Qian Wu,[†] Xin Qi,[†] Qilong Wang,^{††} and Chengwen Song^{†#}

[†]The First Clinical Medical College, Gansu University of Chinese Medicine, Lanzhou 730000, Gansu, China

[‡]Baotou Medical College, Inner Mongolia University of Science & Technology, Baotou, Inner Mongolia, China

[¶]Inner Mongolia Autonomous Region People's Hospital, Hohhot, Inner Mongolia, China

[§]Yangzhou Hospital of Soochow University, Yangzhou, China

^{l0}The Affiliated Hospital of Qingdao University, Shandong, China

[⊥]Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

[#]Panorama Research Institute, Hong Kong, China

^{*}The authors contributed equally.

[†]Corresponding author. Email: Wade200023@outlook.com; wenshao521@panorama-sg.de

Abstract

Background: Short-video platforms are common health-information sources, but lumbar disc herniation (LDH) videos may vary in quality and reliability.

Objective: To compare Chinese-language LDH videos on TikTok and Bilibili and examine associations with platform, source, and engagement.

Methods: This cross-sectional study assessed 200 videos (100 per platform) collected February 28, 2025. Human raters evaluated quality using the 5-point Global Quality Score (GQS) and reliability using an eight-item modified DISCERN scale. Analyses used nonparametric comparisons, Holm-adjusted tests, Spearman correlations, and proportional-odds ordinal regression.

Results: Platforms differed in source, specialty, broad content, and presentation (all $P < 0.001$), but not LDH topic ($P = 0.441$). Physicians produced 85% of TikTok and 61% of Bilibili videos. Median GQS was 3 on both platforms ($P = 0.297$), with similar reliability distributions ($P = 0.816$). Compared with physician videos, independent-user videos had lower adjusted odds of higher GQS (OR 0.29, 95% CI 0.13–0.63; $P = 0.002$) and reliability (OR 0.40, 95% CI 0.18–0.88; $P = 0.023$). Platform and engagement were not independently associated with either outcome.

Conclusions: Source was more informative than platform or engagement. Engagement should not replace clinical credibility when evaluating short-video health information.

Keywords: lumbar disc herniation; health information quality; social media; TikTok; Bilibili; short video; modified DISCERN

1. Introduction

Lumbar disc herniation (LDH) is a common spinal disorder in which displaced disc material can irritate or compress neural structures and produce low-back pain, radiculopathy, sensory or motor



This article is published in *Health Nexus: Digital Health and Medical AI* under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

deficits, and functional limitation (Zhang et al. 2023; Pojskic et al. 2024). Although many patients improve with conservative management, accurate information about red flags, treatment options, indications for surgery, and rehabilitation is important for informed decision-making.

Patients increasingly obtain such information from the Internet and social media rather than relying exclusively on face-to-face clinical encounters. Short-video platforms can improve the accessibility and immediacy of patient education, but the low threshold for content production also allows incomplete, unbalanced, or misleading medical claims to circulate widely (Fareed et al. 2020; He et al. 2023; Sun et al. 2023; Hong et al. 2022). Popularity metrics may further shape visibility without necessarily reflecting scientific quality.

Previous studies have evaluated health-information quality on YouTube and short-video platforms across multiple conditions, including gallstone disease, diabetes, chronic obstructive pulmonary disease, dermatologic disorders, psychiatric conditions, and cancer (Sun et al. 2023; Gad et al. 2022; Yilmaz et al. 2022; Etzel et al. 2021; Arslan et al. 2022; Basch et al. 2021; Comp et al. 2020; Rehman et al. 2022; Yeung et al. 2022; Samuel et al. 2022; Song et al. 2021; Kong et al. 2021; Han et al. 2022; Hu et al. 2022). However, evidence regarding LDH is limited, and few studies have directly compared the content ecology and information quality of TikTok and Bilibili using the same sampling and scoring framework.

Accordingly, this study compared the sources, content characteristics, information quality, and reliability of Chinese-language LDH videos on TikTok and Bilibili. We further examined whether video source and user engagement were associated with higher GQS and modified DISCERN reliability categories after accounting for platform and basic video characteristics.

2. Methods

2.1 Search Strategy and Data Collection

A cross-sectional search was performed on TikTok and Bilibili on February 28, 2025, using the Chinese terms “腰椎间盘突出” (lumbar disc herniation) and “腰椎间盘突出症” (lumbar disc herniation syndrome). For each platform, a new account without prior viewing history was used to reduce personalization. The default platform ranking was recorded, duplicate and ineligible videos were excluded, and 100 videos per platform were retained for analysis (Figure 1).

The search environment was standardized as far as practicable. We did not attempt to reconstruct proprietary recommendation algorithms; instead, videos were sampled according to the default ranking presented to the standardized accounts at the time of data collection.

The final analytic sample comprised 200 videos. Extracted variables included video source, medical specialty, broad content category, disease-knowledge topic, presentation form, likes, comments, saves, shares, days since publication, video duration, and creator follower count. The two platform samples were combined for pooled analyses, with platform retained as a model covariate.

Textual content extraction and AI-assisted scoring using OpenAI GPT-5.6 were conducted for exploratory purposes. Primary analyses used human-rated GQS and modified DISCERN outcomes from the 200-video master dataset; AI-derived scores were not used as primary outcomes.

2.2 Classification of Videos

Video sources were classified as (1) physicians/medical professionals, (2) news organizations, or (3) independent users. Video content was categorized by the predominant presentation type (disease-related knowledge, outpatient/clinical scene, personal experience, or other) and by the principal LDH topic (prevention, diagnosis, symptoms, disease overview, rehabilitation, or treatment). Medical-professional videos were additionally classified by specialty for exploratory subgroup comparisons.

2.3 Video Quality and Reliability Assessments

Information quality was assessed using the Global Quality Score (GQS), a 5-point global rating in which higher scores indicate better flow, completeness, and usefulness. Reliability was evaluated with an eight-item modified DISCERN instrument. In the master dataset, modified DISCERN results were stored as five ordered reliability categories: category 1 corresponded to a score of 1; category 2 to scores 2–3; category 3 to scores 4–5; category 4 to scores 6–7; and category 5 to a score of 8. For the 100 videos with a paired secondary scoring sheet, agreement was evaluated using quadratic-weighted Cohen kappa after applying the same five-category transformation to the secondary modified DISCERN scores.

2.4 Statistical Analysis

Continuous variables were summarized using the median and interquartile range (IQR), and categorical variables using counts and percentages. Platform differences in categorical characteristics were assessed with chi-square statistics; Monte Carlo permutation P values were used for sparse contingency tables. GQS and modified DISCERN categories were compared using Mann-Whitney U tests for two-group comparisons and Kruskal-Wallis tests for three or more groups. Dunn pairwise comparisons were adjusted using the Holm method. Spearman rank correlations were used for video metrics. Because likes, comments, saves, and shares were highly intercorrelated, the primary multivariable model used a single engagement index defined as the mean of standardized $\ln(1+x)$ -transformed likes, comments, saves, and shares, standardized again to one SD. Proportional-odds ordinal logistic regression modeled the odds of being in a higher GQS or modified DISCERN category. The prespecified primary covariates were platform, video source, log-transformed video duration, log-transformed days since upload, and engagement index. Broad content was evaluated separately and in sensitivity analysis because it was strongly structurally associated with platform. Adjusted odds ratios (ORs), 95% confidence intervals (CIs), and two-sided P values are reported; $P < 0.05$ was considered statistically significant.

3. Results

3.1 Video Characteristics

A total of 200 videos were analyzed, comprising 100 TikTok and 100 Bilibili videos. Platform was associated with video source, medical specialty, broad content, and presentation form (all $P < 0.001$), whereas disease-knowledge topic distributions were similar ($P = 0.441$). The largest compositional differences involved broad content and medical specialty: TikTok contained 62 outpatient-scene videos, whereas Bilibili contained none; conversely, disease-knowledge videos accounted for 87% of Bilibili videos compared with 30% of TikTok videos.

Physicians accounted for 85% (85/100) of TikTok videos, followed by independent users (9%) and news organizations (6%). On Bilibili, physicians accounted for 61% (61/100), news organizations for 20%, and independent users for 19%. Broad content also differed markedly: TikTok was dominated by outpatient-scene videos (62%), whereas Bilibili was dominated by disease-knowledge videos (87%) (Figure 2).

3.2 Video Quality and Reliability Assessments

Video source was associated with GQS (Kruskal-Wallis $P = 0.004$) and modified DISCERN reliability category ($P = 0.041$) (Figure 4). After Holm adjustment, independent-user videos had lower GQS than physician videos (adjusted $P = 0.005$) and news-organization videos (adjusted $P = 0.009$), whereas physicians and news organizations did not differ. For modified DISCERN, no individual pairwise contrast remained significant after Holm adjustment despite the significant overall test. Broad content category was not associated with GQS ($P = 0.281$) or modified DISCERN reliability category

($P=0.359$) (Figure 5). Medical-specialty comparisons were also non-significant for GQS ($P=0.694$) and modified DISCERN ($P=0.619$); these sparse subgroup analyses are best considered exploratory.

In the available paired 100-video scoring subset, exact agreement was 66% for GQS and 79% for the five-level modified DISCERN reliability category. Quadratic-weighted kappa was 0.556 (95% bootstrap CI 0.368–0.717) for GQS and 0.673 (95% bootstrap CI 0.497–0.827) for modified DISCERN category, indicating moderate-to-substantial ordinal agreement. Secondary ratings were available for 100 videos; therefore, these agreement estimates apply only to that subset.

3.3 Correlation and Ordinal Regression Analyses

Engagement variables were strongly intercorrelated (Table 8). The strongest association was between saves and shares (Spearman $\rho=0.9214$, $P<0.001$). Likes were not significantly correlated with duration ($\rho=-0.0959$, $P=0.177$). Engagement metrics were not significantly correlated with human-rated quality in the full 200-video dataset (Table 9). Likes, comments, saves, and shares were not significantly correlated with GQS or modified DISCERN category. Only follower count showed a weak unadjusted positive correlation with GQS ($\rho=0.144$, $P=0.042$), while its correlation with modified DISCERN category did not reach significance ($\rho=0.130$, $P=0.067$).

4. Discussion

4.1 Principal Findings

Using the verified 200-video master dataset, this study yielded four clinically relevant findings. First, TikTok and Bilibili showed markedly different content ecologies, particularly in uploader composition, medical specialty, broad content category, and presentation form. Second, despite these compositional differences, the overall quality of LDH-related videos remained moderate, and platform itself was not independently associated with a higher GQS or modified DISCERN category. Third, information provenance was more informative than platform identity: compared with physician-generated videos, independent-user videos had 71% lower adjusted odds of occupying a higher GQS category and 60% lower adjusted odds of occupying a higher modified DISCERN category. Fourth, popularity was not a reliable marker of quality. Likes, comments, saves, and shares were strongly intercorrelated, but neither individual engagement metrics nor the composite engagement index was independently associated with human-rated quality or reliability. Taken together, these findings suggest that who produces the information may matter more than where it is posted or how much attention it receives.

4.2 Comparison With Existing LDH and Spine-Related Evidence

Our findings are broadly consistent with the emerging literature on online information for spinal disorders. A recent four-platform study of 384 LDH-related short videos from Douyin, Bilibili, Kuaishou, and Xiaohongshu reported generally low information quality, with a mean GQS of approximately 2 and low DISCERN scores; treatment was the most frequently covered topic, appearing in more than four-fifths of videos (Wang *et al.* 2025). The median GQS of 3 observed in our sample therefore appears somewhat higher, but direct numerical comparison should be avoided because the studies differed in platform mix, sampling frame, ranking strategy, and implementation of DISCERN. More importantly, both studies converge on the same practical message: highly visible LDH videos frequently provide only moderate or incomplete patient education and should not be treated as substitutes for individualized clinical assessment. Earlier work on disc-herniation videos on YouTube similarly found low overall quality and no significant relationship between video popularity and DISCERN or JAMA scores (Gokcen and Gumussuyu 2019).

Comparable patterns have also been reported for other spine and musculoskeletal conditions. In scoliosis-related TikTok videos, professional creators achieved higher quality scores, whereas patient identity was independently associated with lower GQS in an ordinal regression model (Zhou *et*

al. 2025). A 2026 study of cervical-spondylosis videos found an overall median GQS and modified DISCERN score of 3, with higher scores for orthopedic and rehabilitation professionals and only weak-to-moderate relationships between engagement and quality (Lin et al. 2026). These findings are closely aligned with our observation that LDH videos occupy a middle rather than high quality range and that professional provenance is an important discriminator of credibility.

4.3 Platform Ecology Does Not Necessarily Equal Platform Quality

The marked differences in content composition between TikTok and Bilibili are important, but our multivariable analysis suggests that platform identity should not be interpreted as a direct quality determinant. This distinction is supported by recent cross-platform musculoskeletal studies showing that the direction of platform effects is disease-specific. For osteoarthritis, TikTok achieved higher GQS, JAMA, and HONcode scores than Bilibili, whereas modified DISCERN did not differ significantly (Zuo et al. 2025). In knee-arthroplasty videos, TikTok showed greater engagement and higher scores across several quality instruments, while Bilibili offered more structured educational formats and more diverse professional backgrounds (Wei et al. 2026). By contrast, an ankylosing-spondylitis study found Bilibili superior in academic rigor, although both platforms remained limited in practical utility (Liu et al. 2026). The lack of a consistent platform hierarchy across conditions suggests that video quality emerges from the interaction of uploader mix, topic, presentation conventions, sampling strategy, and the evaluation instrument rather than from a stable intrinsic advantage of one platform over another.

This interpretation is especially relevant to our dataset because TikTok contained more physician-generated and outpatient-scene videos, whereas Bilibili contained more disease-knowledge videos and a larger contribution from news organizations and independent users. Such ecological differences can shape crude platform comparisons even when the platform coefficient becomes non-significant after adjustment. Accordingly, future comparative studies should report both platform-level descriptive differences and source-adjusted models rather than treating platform as a single causal exposure.

4.4 Professional Provenance and the Limits of Credential-Based Trust

The most robust signal in our study was uploader source. Independent-user videos were significantly less likely than physician-generated videos to achieve higher quality and reliability categories, and this pattern remained after adjustment for platform, content, duration, recency, and engagement. This finding is consistent with recent studies of scoliosis, osteoarthritis, knee arthroplasty, ankylosing spondylitis, and cervical spondylosis, all of which identified professional involvement as an important contributor to information quality in at least some domains (Zhou et al. 2025; Zuo et al. 2025; Wei et al. 2026; Liu et al. 2026; Lin et al. 2026). It also supports the practical value of transparent creator identity and credential verification on health-information platforms.

However, professional authorship should not be equated with automatically high-quality communication. In our study, median scores among the major source groups remained near the middle of the scales, indicating that credentialed creators may still omit references, areas of uncertainty, alternatives, prognosis, or limitations of treatment. Short-form communication can compress complex clinical decisions into simplified messages, and professional status cannot compensate for absent evidence or incomplete context. The more defensible implication is therefore not simply to promote physician-generated videos, but to encourage physician-generated videos that also provide source attribution, balanced discussion of options, explicit uncertainty, and clear boundaries between general education and individualized medical advice.

4.5 Popularity Is an Attention Signal, Not a Quality Signal

Engagement measures were highly correlated with one another, but individual metrics were not significantly associated with human-rated GQS or modified DISCERN reliability categories. The

adjusted engagement index was also not significantly associated with either outcome. This result is consistent with the earlier YouTube disc-herniation study (Gokcen and Gumussuyu 2019) and with the recent ankylosing-spondylitis analysis, in which quality and popularity were essentially disconnected (Liu *et al.* 2026).

At the same time, the broader literature is heterogeneous. Osteoarthritis-related videos showed positive associations between several engagement metrics and quality measures (Zuo *et al.* 2025), while scoliosis videos identified collections and duration as independent correlates of higher GQS (Zhou *et al.* 2025). Knee-arthroplasty research also found that engagement correlated with some production- or usability-oriented measures but not consistently with modified DISCERN (Wei *et al.* 2026), and cervical-spondylosis videos showed only weak-to-moderate correlations (Lin *et al.* 2026). These divergent results are important because they argue against a universal rule that popular videos are either better or worse. Engagement is likely influenced by topic salience, emotional appeal, creator visibility, video length, production quality, and platform-specific exposure. Therefore, likes or saves should be treated as measures of reach or attention, not as surrogate markers of clinical accuracy.

4.6 Content Coverage and Clinical Relevance

Broad content category was not significantly associated with GQS or modified DISCERN in our analyses, but this does not mean that all content domains were equally complete or clinically useful. The recent four-platform LDH study showed a pronounced emphasis on treatment, with much less coverage of diagnosis and other aspects of disease management (Wang *et al.* 2025). For LDH, patient education is most useful when it places treatment within a broader clinical pathway that includes symptom interpretation, red-flag recognition, expected natural history, conservative management, indications for imaging, indications for surgery, rehabilitation, and the limits of self-management. The WFNS recommendations similarly emphasize that diagnosis and management require integration of clinical and radiologic information rather than reliance on isolated symptoms or imaging findings (Pojskic *et al.* 2024).

This distinction also highlights a limitation of broad global scoring tools. A video may receive a moderate GQS because it is understandable and reasonably useful yet still omit a clinically important domain. Conversely, a focused video on one narrow topic may be accurate but score lower because it is intentionally incomplete. Future LDH studies would therefore benefit from pairing GQS and modified DISCERN with a prespecified, guideline-based content checklist that separately evaluates core domains such as diagnosis, red flags, conservative care, rehabilitation, surgical indications, prognosis, and safety. This would allow investigators to distinguish general communication quality from disease-specific completeness.

4.7 Measurement and Methodological Implications

GQS and modified DISCERN remain useful because they are simple, reproducible, and widely used in health-information research, but they assess overlapping rather than identical constructs. GQS emphasizes overall flow, usefulness, and patient-oriented quality, whereas modified DISCERN emphasizes transparency and reliability. Recent musculoskeletal studies increasingly combine these measures with JAMA benchmarks, PEMAT, VIQI, or HONcode to capture source transparency, audiovisual quality, understandability, and actionability (Zuo *et al.* 2025; Wei *et al.* 2026; Lin *et al.* 2026). This multidimensional approach may be particularly valuable for LDH, because visually demonstrated exercises or self-management maneuvers can be misleading even when the accompanying text appears credible.

Our study also illustrates why ordinal modeling is preferable to treating these bounded scores as count outcomes. Both GQS and the retained modified DISCERN categories are ordered measures with a finite number of levels. Proportional-odds ordinal regression therefore provides an interpretable

estimate of the odds of occupying a higher quality category while respecting the ordinal structure. Nevertheless, the master dataset retained modified DISCERN as five ordered categories rather than the original item-level or exact total score, which reduced granularity and precluded item-level reliability analysis across all 200 videos. Future studies should preserve the original item-level ratings whenever possible.

4.8 Implications for Patients, Clinicians, and Platforms

The findings have implications at three levels. For patients, high engagement should not be interpreted as evidence of medical reliability; creator credentials, cited sources, balanced discussion, and consistency with professional guidance are more defensible signals. For clinicians, the goal should be to combine professional credibility with communication strategies suited to short-form media: concise but structured explanations, clear red flags, realistic expectations, references to authoritative guidance, and explicit advice about when in-person evaluation is necessary. For platforms, verification badges alone are unlikely to be sufficient. Health-related recommendation and moderation systems could place greater weight on transparent professional identity, source citation, disclosure of uncertainty, and removal or labeling of demonstrably misleading claims rather than relying primarily on interaction volume.

Importantly, our data do not directly test platform algorithms, ranking mechanisms, or downstream patient behavior. These recommendations should therefore be understood as risk-reduction strategies derived from observed associations rather than evidence that a particular algorithm causes misinformation. Prospective studies linking video exposure to patient knowledge, treatment choices, adherence, and clinical outcomes would provide a stronger basis for platform-level policy.

4.9 Strengths and Limitations

This study has several strengths. Videos were collected using a standardized same-day search strategy with newly created accounts, the two platforms contributed equal sample sizes, and the final analysis was based on a verified 200-row master dataset. Human ratings were separated from engagement metrics, nonparametric methods were used for ordinal outcomes, pairwise testing incorporated multiplicity control, and the multivariable analysis used proportional-odds ordinal regression rather than Poisson regression. In addition, the inclusion of platform, source, content, duration, recency, and a standardized engagement index allowed the analysis to distinguish compositional platform differences from independent predictors of quality.

Several limitations should be considered. First, the study represents a cross-sectional snapshot of highly ranked videos on a single date; platform rankings, creator behavior, and engagement counts are dynamic. Second, only the top 100 eligible videos per platform were included, so the sample reflects high-visibility search results rather than the full universe of LDH content. Third, the two platforms have different user communities, video conventions, and exposure mechanisms, and our dataset did not contain denominator-level impressions, watch time, completion rate, or recommendation history; consequently, engagement counts cannot be interpreted as standardized exposure measures. Fourth, modified DISCERN was retained as five ordered reliability categories rather than the original item-level or exact total score, limiting granularity. Fifth, paired secondary ratings were available for only 100 videos, so the inter-rater reliability estimates apply to that subset. Sixth, some specialty subgroups were extremely small, including only one neurosurgery video, and several specialties were absent from one platform, limiting power for specialty comparisons. Seventh, GQS and modified DISCERN do not directly measure factual accuracy of visual demonstrations, guideline concordance, readability, or patient actionability. Finally, because this was an observational content analysis, the associations between uploader source and quality should not be interpreted as causal.

5. Conclusions

Among 200 Chinese-language LDH videos, TikTok and Bilibili differed markedly in source and content composition but not in overall GQS or modified DISCERN reliability distributions. Independent-user source was independently associated with lower odds of higher quality and reliability, whereas the engagement index was not. These findings support prioritizing transparent professional provenance and evidence-based content standards rather than popularity metrics when patients use short-video platforms for LDH information.

Declarations

Author Contributions

Conceptualization: Qilong Wang and Chengwen Song. **Methodology:** Chen Wang, Chao Yan, Xin Qi, and Qian Wu. **Formal Analysis:** Chen Wang, Chao Yan, Xin Qi, and Qian Wu. **Investigation and Data Curation:** Chen Wang, Chao Yan, Yibo Ma, Guoyi Wang, Yuyu Zhang, and Yayan Wang. **Writing – Original Draft:** Chen Wang and Chao Yan. **Writing – Review and Editing:** Qilong Wang, Chengwen Song, Xin Qi, and Qian Wu. **Supervision, Project Administration, and Funding Acquisition:** Qilong Wang and Chengwen Song. All authors contributed to interpretation of the results, critically reviewed and approved the final version of the manuscript. Chen Wang and Chao Yan contributed equally and are co-first authors.

Funding

This study was supported by the Natural Science Foundation of Gansu Province (No. 23JRRA1280), the Gansu Provincial People's Hospital National Research Project Cultivation Program General Project (No. 19SYPYB-29), the Gansu University of Chinese Medicine Graduate Student Innovation and Entrepreneurship Fund Project (No. 2025CXCY-085), and Panorama Research Institute (Sponsorship No. PSG-PRI-SPN-2026-000019).

Competing Interests

The authors declare that they have no competing financial or non-financial interests relevant to the work reported in this article.

Acknowledgements

The authors thank the study team members who assisted with video screening, scoring, and data verification.

Ethics Approval and Consent to Participate

The Ethics Committee of Gansu Provincial People's Hospital determined that ethics approval was not required because the study analyzed publicly available online video content and did not involve intervention, interaction, or collection of identifiable private information from human participants. Informed consent to participate was not applicable.

Consent for Publication

Not applicable. The study did not collect identifiable private information from human participants.

Clinical-Trial Registration

Not applicable.

Data Availability

The data supporting the findings of this study are available from the author upon reasonable request, subject to applicable platform terms and data-sharing restrictions.

Generative AI and AI-Assisted Technology Declaration

OpenAI GPT-5.6 was used for exploratory textual content extraction and AI-assisted scoring. AI-derived scores were not used as primary outcomes.

References

1. Arslan, Can, Elif Ceren Aksahin, Rahime Burcu Nur Yilmaz, and Derya Germec Cakan. 2022. "Does YouTube TM Offer High-Quality Information About Nasoalveolar Molding?" *The Cleft Palate Craniofacial Journal* 61, no. 1: 5–11. <https://doi.org/10.1177/10556656221115025>
2. Balci, Ali Safa, Kubra Şerefoglu Cabuk, Husna Topcu, Ayse Cetin Efe, and Mehmet Goksel Ulas. 2022. "Evaluation of the reliability, utility, and quality of the lid loading videos on YouTube." *International Ophthalmology* 43, no. 6: 2065–2072. <https://doi.org/10.1007/s10792-022-02606-w>
3. Basch, Corey H, Joseph Fera, Isabela Pierce, and Charles E Basch. 2021. "Promoting Mask Use on TikTok: Descriptive, Cross-sectional Study." *JMIR Public Health and Surveillance* 7, no. 2: e26392. <https://doi.org/10.2196/26392>
4. Bernard, André, Morgan Langille, Stephanie Hughes, Caren Rose, Desmond Leddin, and Sander Veldhuyzen van Zanten. 2007. "A Systematic Review of Patient Inflammatory Bowel Disease Information Resources on the World Wide Web." *The American Journal of Gastroenterology* 102, no. 9: 2070–2077. <https://doi.org/10.1111/j.1572-0241.2007.01325.x>
5. Charnock, D, S Shepperd, G Needham, and R Gann. 1999. "DISCERN: an instrument for judging the quality of written consumer health information on treatment choices." *Journal of Epidemiology and Community Health* 53, no. 2: 105–111. <https://doi.org/10.1136/jech.53.2.105>
6. Comp, Geoffrey, Sean Dyer, and Michael Gottlieb. 2020. "Is TikTok The Next Social Media Frontier for Medicine?" *AEM Education and Training* 5, no. 3. <https://doi.org/10.1002/aet2.10532>
7. Yiming Dou, Yiming Zhang, Yang Liu, Xun Sun, Xinyu Liu, Bin Li, and Qiang Yang. 2025. "Role of macrophage in intervertebral disc degeneration." *Bone Research* 13, no. 1: 15–15. <https://doi.org/10.1038/s41413-024-00397-7>
8. Etzel, Christine M, Steven L Bokshan, Timothy A Forster, and Brett D Owens. 2021. "A quality assessment of YouTube content on shoulder instability." *The Physician and Sportsmedicine* 50, no. 4: 289–294. <https://doi.org/10.1080/00913847.2021.1942286>
9. Fareed, Naleef, Christine M. Swoboda, Pallavi Jonnalagadda, and Timothy R. Huerta. 2020. "Persistent digital divide in health-related internet use among cancer survivors: findings from the Health Information National Trends Survey, 2003–2018." *Journal of Cancer Survivorship* 15, no. 1: 87–98. <https://doi.org/10.1007/s11764-020-00913-8>
10. Ferhatoglu, Murat Ferhat, Abdulcabbar Kartal, Ugur Ekici, and Alp Gurkan. 2019. "Evaluation of the Reliability, Utility, and Quality of the Information in Sleeve Gastrectomy Videos Shared on Open Access Video Sharing Platform YouTube." *Obesity Surgery* 29, no. 5: 1477–1484. <https://doi.org/10.1007/s11695-019-03738-2>
11. Francisco, Vera, Jesús Pino, Miguel Ángel González-Gay, Francisca Lago, Jaro Karppinen, Osmo Tervonen, Ali Mobasheri, and Oreste Gualillo. 2021. "A new immunometabolic perspective of intervertebral disc degeneration." *Nature Reviews Rheumatology* 18, no. 1: 47–60. <https://doi.org/10.1038/s41584-021-00713-z>
12. Gad, Bishoy, Srinidhi Shanmugasundaram, Abhishek Kumar, and Pratik Shukla. 2022. "Quality and Reliability of YouTube Videos on Uterine Fibroid Embolization." *Journal of the American College of Radiology* 19, no. 7: 905–912. <https://doi.org/10.1016/j.jacr.2022.03.014>

13. Gokcen, H. Bahadir, and Gurkan Gumussuyu. 2019. "A Quality Analysis of Disc Herniation Videos on YouTube." *World Neurosurgery* 124: e799–e804. <https://doi.org/10.1016/j.wneu.2019.01.146>
14. Han, Jiajun, Yifan Shi, and Haitao Ma. 2022. "Assessment of videos related to lung nodules in China." *Frontiers in Surgery* 9. <https://doi.org/10.3389/fsurg.2022.1019212>
15. He, Zixuan, Zhijie Wang, Yihang Song, Yilong Liu, Le Kang, Xue Fang, Tongchang Wang, Xuanming Fan, Zhaoshen Li, Shuling Wang, and Yu Bai. 2023. "The Reliability and Quality of Short Videos as a Source of Dietary Guidance for Inflammatory Bowel Disease: Cross-sectional Study." *Journal of Medical Internet Research* 25: e41518–e41518. <https://doi.org/10.2196/41518>
16. Hong, Y. Alicia, Soo Yee, Pramita Bagchi, Hee-soon Juon, Sojung Claire Kim, and Daisy Le. 2022. "Social media-based intervention to promote HBV screening and liver cancer prevention among Korean Americans: Results of a pilot study." *DIGITAL HEALTH* 8: 205520762210762. <https://doi.org/10.1177/20552076221076257>
17. Hu, Ren-hao, Hai-bin Zhang, Biao Yuan, Ke-hui Zhang, Jia-yi Xu, Xi-mao Cui, Tao Du, Chun Song, Shun Zhang, and Xiao-hua Jiang. 2022. "Quality and accuracy of gastric cancer related videos in social media videos platforms." *BMC Public Health* 22, no. 1. <https://doi.org/10.1186/s12889-022-14417-w>
18. Huang, Yeheng, Jing Sun, Sunlong Li, Yifeng Shi, Lianggao Yu, Aimin Wu, and Xiangyang Wang. 2024. "Isoliquiritigenin mitigates intervertebral disc degeneration induced by oxidative stress and mitochondrial impairment through a PPAR γ -dependent pathway." *Free Radical Biology and Medicine* 225: 98–111. <https://doi.org/10.1016/j.freeradbiomed.2024.10.001>
19. Dana Kerr, Wenyan Zhao, and Jon D. Lurie. 2014. "What Are Long-term Predictors of Outcomes for Lumbar Disc Herniation? A Randomized and Observational Study." *Clinical Orthopaedics and Related Research* 473, no. 6: 1920–1930. <https://doi.org/10.1007/s11999-014-3803-7>
20. Kong, Wenwen, Shijie Song, Yuxiang Chris Zhao, Qinghua Zhu, and Ling Sha. 2021. "TikTok as a Health Information Source: Assessment of the Quality of Information in Diabetes-Related Videos." *Journal of Medical Internet Research* 23, no. 9: e30409. <https://doi.org/10.2196/30409>
21. Li, Fudong, Shaofei Li, Yangyang Shi, Feng Lin, Lining Rui, Jiangang Shi, and Kaiqiang Sun. 2024. "Glutathione: A Key Regulator of Extracellular Matrix and Cell Death in Intervertebral Disc Degeneration." *Mediators of Inflammation* 2024, no. 1. <https://doi.org/10.1155/2024/4482642>
22. Lin, Jie-Min, Kun-Hui Li, and Cheng Lin. 2026. "The quality and reliability of short videos about cervical spondylosis on TikTok and Redbook: a cross-sectional study." *BMC Musculoskeletal Disorders* 27, no. 1. <https://doi.org/10.1186/s12891-026-10105-7>
23. Liu, Ruofan, Xinyang Wang, Taotao Wang, Yunfei Li, Shuangshuang Shang, and Chuanbing Huang. 2026. "Professional Creators Provide Higher-Quality Ankylosing Spondylitis Content on TikTok and Bilibili: A Cross-Sectional Study." *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 63. <https://doi.org/10.1177/00469580261436338>
24. Mueller, Simon M, Valentina N S Hongler, Pierre Jungo, Lucian Cajacob, Simon Schwegler, Esther H Steveling, Zita-Rose Manjaly Thomas, Oliver Fuchs, Alexander Navarini, Kathrin Scherer, and Oliver Brandt. 2020. "Fiction, Falsehoods, and Few Facts: Cross-Sectional Study on the Content-Related Quality of Atopic Eczema-Related Videos on YouTube." *Journal of Medical Internet Research* 22, no. 4: e15599. <https://doi.org/10.2196/15599>
25. Novais, Emanuel J., Victoria A. Tran, Shira N. Johnston, Kayla R. Darris, Alex J. Roupas, Garrett A. Sessions, Irving M. Shapiro, Brian O. Diekman, and Makarand V. Risbud. 2021. "Long-term treatment with senolytic drugs Dasatinib and Quercetin ameliorates age-dependent intervertebral disc degeneration in mice." *Nature Communications* 12, no. 1. <https://doi.org/10.1038/s41467-021-25453-2>
26. Pojskic, Mirza, Erica Bisson, Joachim Oertel, Toshihiro Takami, Corinna Zygourakis, and Francesco Costa. 2024. "Lumbar disc herniation: Epidemiology, clinical and radiologic diagnosis

- WFNS spine committee recommendations." *World Neurosurgery*: X 22: 100279–100279. <https://doi.org/10.1016/j.wnsx.2024.100279>
27. Qi, Xin, Qi Lu, Bo Li, Sha Liu, Lei Zhang, Runwu Xiang, Qilong Wang, and Dongrong Zhao. 2025. "Videos in short video sharing platforms as a source of information on bipolar disorder: a cross-sectional content analysis study." *Frontiers in Public Health* 13. <https://doi.org/10.3389/fpubh.2025.1627885>
28. Rehman, Rafey, Sazid Hasan, Hassan Akram, and Marla Jahnke. 2022. "TikTok as a Source of Dermatologic Information on Atopic Dermatitis." *Dermatitis* 33, no. 6S: S133–S134. <https://doi.org/10.1097/der.0000000000000830>
29. Samuel, L., H. Tang, and C.H. Basch. 2022. "TikTok: a far-reaching opportunity for health professionals to address weight loss." *Public Health* 213: 94–99. <https://doi.org/10.1016/j.puhe.2022.10.010>
30. Song, Shijie, Xiang Xue, Yuxiang Chris Zhao, Jinhao Li, Qinghua Zhu, and Mingming Zhao. 2021. "Short-Video Apps as a Health Information Source for Chronic Obstructive Pulmonary Disease: Information Quality Assessment of TikTok Videos." *Journal of Medical Internet Research* 23, no. 12: e28318. <https://doi.org/10.2196/28318>
31. Sun, Fei, Shusen Zheng, and Jian Wu. 2023. "Quality of Information in Gallstone Disease Videos on TikTok: Cross-sectional Study." *Journal of Medical Internet Research* 25: e39162–e39162. <https://doi.org/10.2196/39162>
32. Bing Tan, Shanlin Xiang, Yuhao Zheng, Jianyuan Ouyang, and Nian Zhou. 2025. "Association of dyslipidemia with intervertebral disc degeneration: a case-control study." *European journal of medical research* 30, no. 1: 194–194. <https://doi.org/10.1186/s40001-025-02455-0>
33. Tang, Yingchuang, Kai Zhang, Hongyou Zhou, Chenchen Zhang, Zixiang Liu, Hao Chen, Hanwen Li, and Kangwu Chen. 2023. "Transplantation of active nucleus pulposus cells with a keep-charging hydrogel microsphere system to rescue intervertebral disc degeneration." *Journal of Nanobiotechnology* 21, no. 1. <https://doi.org/10.1186/s12951-023-02226-1>
34. Trefilova, Vera V., Natalia A. Shnayder, Marina M. Petrova, Daria S. Kaskaeva, Olga V. Tutynina, Kirill V. Petrov, Tatiana E. Popova, Olga V. Balberova, German V. Medvedev, and Regina F. Nasyrova. 2021. "The Role of Polymorphisms in Collagen-Encoding Genes in Intervertebral Disc Degeneration." *Biomolecules* 11, no. 9: 1279–1279. <https://doi.org/10.3390/biom11091279>
35. 2019. "A Brief Review of the Degenerative Intervertebral Disc Disease." *Medical Archives*. <https://doi.org/10.5455/medarh.2019.73.421-424>
36. Wang, Yu, Ye Wu, Bo Zhang, Cheng Zheng, Cheng Hu, Chuan Guo, Qingquan Kong, and Yunbing Wang. 2023. "Repair of degenerative nucleus pulposus by polyphenol nanosphere-encapsulated hydrogel gene delivery system." *Biomaterials* 298: 122132–122132. <https://doi.org/10.1016/j.biomaterials.2023.122132>
37. Ze Wang, Jiawei Ma, Yun Sun, Zebin Jin, Rukang Zheng, Yuanyuan Li, Heng Yu, Haobo Ye, Yaosen Wu, Xinjiang Ge, and Zexin Chen. 2024. "Isorhapontigenin delays senescence and matrix degradation of nucleus pulposus cells via PI3K/AKT/mTOR-mediated autophagy pathway in vitro and alleviates intervertebral disc degeneration in vivo." *International Immunopharmacology* 139: 112717–112717. <https://doi.org/10.1016/j.intimp.2024.112717>
38. Wang, Yuehua, Sheng Lian, Longteng Liao, and Wenyu Wang. 2025. "Assessment of information quality and reliability of short videos related to lumbar disc herniation on selected video platforms." *DIGITAL HEALTH* 11. <https://doi.org/10.1177/20552076251393278>
39. Wang, Qilong, Siheng Ma, Dongmei Ma, Chen Wang, Xin Qi, Sha Liu, Lei Zhang, Runwu Xiang, and Dongrong Zhao. 2026. "Video in Chinese short video sharing platforms as a source of information on sleep disorders: A cross-sectional content analysis study." *DIGITAL HEALTH* 12. <https://doi.org/10.1177/20552076261415944>
40. Wei, Bingqi, Xinyue Zhang, Liu Chen, Xingyue Ren, Yijing Li, Feiyang Chen, Luwei Zhang,

- Yunfang Fan, Zimeng Xie, Jiarong Li, Yueyang Chen, and Shangzeng Wang. 2026. "A comparative study of medical information quality and dissemination efficacy of knee arthroplasty videos in Bilibili/TikTok short video platforms." *DIGITAL HEALTH* 12. <https://doi.org/10.1177/20552076261421072>
41. Winkler, J.R., and J.B. Davies. 1986. "Accurate Field Analysis of the Propagation of Elastic Waves Through an Acoustic Microscope, Part I: Theory." *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 33, no. 6: 644–656. <https://doi.org/10.1109/t-uffc.1986.26879>
 42. Wu, Shaobo, Jin Wang, Minglin Wang, Kaisheng Zhou, Dageng Huang, Yilei Zhang, and Haihong Zhang. 2024. "Glucose deprivation-induced disulfidptosis in human nucleus pulposus cells: a novel pathological mechanism of intervertebral disc degeneration." *Biology Direct* 19, no. 1. <https://doi.org/10.1186/s13062-024-00528-4>
 43. Xue, Xiaoqiang, Xinyi Yang, Weifeng Xu, Guanghua Liu, Yi Xie, and Zhigang Ji. 2022. "TikTok as an Information Hodgepodge: Evaluation of the Quality and Reliability of Genitourinary Cancers Related Content." *Frontiers in Oncology* 12. <https://doi.org/10.3389/fonc.2022.789956>
 44. Yang, Shaofeng, Linghui Li, Liguozhu, Chao Zhang, Zhaoyong Li, Yantao Guo, Ying Nie, and Zhenhua Luo. 2019. "Bu-Shen-Huo-Xue-Fang modulates nucleus pulposus cell proliferation and extracellular matrix remodeling in intervertebral disk degeneration through miR-483 regulation of Wnt pathway." *Journal of Cellular Biochemistry* 120, no. 12: 19318–19329. <https://doi.org/10.1002/jcb.26760>
 45. Yang, Shijie, Jing Zhan, and Xiequn Xu. 2023. "Is TikTok a high-quality source of information on thyroid cancer?" *Endocrine* 81, no. 2: 270–276. <https://doi.org/10.1007/s12020-023-03332-8>
 46. Hantao Yang, Xiaoyan Chen, Jun Chen, Yansong Dong, Ya-Fang Huang, Qin L, Jie Tan, and Weihong Yi. 2024. "The pathogenesis and targeted therapies of intervertebral disc degeneration induced by cartilage endplate inflammation." *Frontiers in Cell and Developmental Biology* 12: 1492870–1492870. <https://doi.org/10.3389/fcell.2024.1492870>
 47. Yeung, Anthony, Enoch Ng, and Elia Abi-Jaoude. 2022. "TikTok and Attention-Deficit/Hyperactivity Disorder: A Cross-Sectional Study of Social Media Content Quality." *The Canadian Journal of Psychiatry* 67, no. 12: 899–906. <https://doi.org/10.1177/07067437221082854>
 48. Yilmaz, Mehmet, Yusuf Sahin, Ibrahim Hacibey, Salih Zeki Sonmez, and Ahmet Yaser Muslumanoglu. 2022. "Quality and utility of YouTube videos about mPCNL." *Urolithiasis* 51, no. 1. <https://doi.org/10.1007/s00240-022-01374-7>
 49. Zhang, Andrew S., Andrew Xu, Kashif Ansari, Kyle Hardacker, George Anderson, Daniel Alsoof, and Alan H. Daniels. 2023. "Lumbar Disc Herniation: Diagnosis and Management." *The American Journal of Medicine* 136, no. 7: 645–651. <https://doi.org/10.1016/j.amjmed.2023.03.024>
 50. Zhou, Minhao, Alekos A. Theologis, and Grace D. O'Connell. 2023. "Understanding the etiopathogenesis of lumbar intervertebral disc herniation: From clinical evidence to basic scientific research." *JOR SPINE* 7, no. 1. <https://doi.org/10.1002/jsp2.1289>
 51. Zhou, Qinxin, Jianzeng Shen, Dongdong Cao, Weijie Yu, and Jixin Chen. 2025. "Quality and accuracy of scoliosis-related short videos on TikTok reviewed in China in August 2025." *Scientific Reports* 15, no. 1. <https://doi.org/10.1038/s41598-025-27684-5>
 52. Zuo, Qi-Heng, Kai Du, Ao Li, Chen-Yu Zhang, Ren Guo, Ping Chen, Wei-Shuai Du, Yong-Li Zuo, and Shu-Ming Li. 2025. "Quality and reliability of osteoarthritis-related health information on short video platforms: a cross-sectional comparative study of TikTok and Bilibili." *Frontiers in Digital Health* 7. <https://doi.org/10.3389/fdgh.2025.1623247>

Table 1. Global Quality Score criteria used to assess the quality of lumbar disc herniation-related videos.

Scale	Description
1	poor quality, poor flow of the site, most information missing, not at all useful for patients
2	Generally poor quality and poor flow, some information listed but many important topics missing, of very limited use to patients
3	Moderate quality and flow; some important information is adequately discussed while other topics are incomplete; somewhat useful for patients.
4	Good quality and generally good flow, most of the relevant information is listed, but some topics not covered, useful for patients
5	Excellent quality and excellent flow, very useful for patients

Table 2. Reliability-related DISCERN questions used in the eight-item modified DISCERN assessment.

DISCERN	Description
Question 1	Are the aims clear?
Question 2	Does it achieve its aims?
Question 3	Is it relevant?
Question 4	Is it clear what sources of information were used to compile the publication (other than the author or producer)?
Question 5	Is it clear when the information used or reported in the publication was produced?
Question 6	Is it balanced and unbiased?
Question 7	Does it provide details of additional sources of support and information?
Question 8	Does it refer to areas of uncertainty?

Table 3. Operational interpretation of GQS and modified DISCERN scores.

Scale, score	Interpretation
GQS	
1	Very poor
2	Poor
3	Fair/moderate
4	Good
5	Excellent
Modified DISCERN	
1	Unreliable
2-3	Less reliable
4-5	Fairly reliable
6-7	Relatively reliable
8	Reliable

Table 4. Characteristics of lumbar disc herniation-related videos on TikTok and Bilibili.

Variable	Total (N=200)	TikTok (n=100)	Bilibili (n=100)	P value
Video source				<0.001
Physicians	146	85	61	
News organizations	26	6	20	
Independent users	28	9	19	
Medical specialty				<0.001
Orthopedic surgery	81	58	23	
Pain medicine	19	10	9	
Rehabilitation	22	6	16	
Neurosurgery	1	1	0	
Chinese medicine	19	5	14	
Imaging department	43	4	39	
Other	15	15	0	
Broad content category				<0.001
Disease-related knowledge	117	30	87	
Outpatient scene	62	62	0	
Personal experience	19	8	11	
Other	2	0	2	
LDH topic				0.441
Prevention	34	18	16	
Diagnosis	29	17	12	
Symptoms	29	14	15	
Disease overview	26	10	16	
Rehabilitation	43	18	25	
Treatment	39	23	16	
Presentation format				<0.001
Expert monologue	134	75	59	
Animation	20	11	9	
Patient vlog	29	1	28	
Images and literature	16	12	4	
Other	1	1	0	

Table 5. Engagement characteristics of TikTok videos according to source and content category.

Variable	Likes	Comments	Shares	Saves	Days since published	Duration
Video sources (n=100), median (IQR)						
Physicians (n=85)	10876 (5, 320000)	402 (0, 13000)	3383 (0, 146000)	4603 (2, 393000)	35 (1, 605)	67 (6, 218)
News organizations (n=6)	1663.5 (33, 16000)	513 (2, 2682)	689.5 (0, 2268)	1836 (34, 3908)	354.25 (1, 782)	123.75 (21, 721)
Independent users (n=9)	19806 (1687, 44000)	431 (93, 8302)	7756 (48, 22000)	16766 (277, 43000)	23 (4, 66)	95 (11, 542)
Video content (n=100), median (IQR)						
Disease-related knowledge (n=30)	6878.75 (27, 320000)	378.25 (2, 13000)	3433 (0, 110000)	2331.5 (8, 134000)	61.75 (1, 782)	130.25 (11, 542)
Outpatient scene (n=62)	10894.5 (5, 164000)	395.5 (0, 6660)	2809 (0, 146000)	4696.75 (2, 393000)	27.5 (1, 142)	61.25 (6, 218)
Personal experience (n=8)	22293 (449, 44000)	583.5 (3, 8302)	4960.5 (204, 22000)	17417 (765, 43000)	25.25 (2, 66)	168.25 (32, 721)

Table 6. Engagement characteristics of Bilibili videos according to source and content category.

Variable	Likes	Comments	Shares	Saves	Days since published	Duration
Video sources (n=100), median (IQR)						
Physicians (n=61)	1670 (14, 77000)	183 (1, 1927)	1236 (1, 29000)	2737 (10, 64000)	151 (2, 539)	202 (14, 3587)
News organizations (n=20)	3465.5 (135, 46000)	427 (6, 3308)	2815 (74, 44000)	9134 (273, 64000)	249.75 (11, 435)	858.75 (51, 7545)
Independent users (n=19)	2697.5 (39, 32000)	371.5 (0, 2586)	1465.5 (1, 19000)	7258.5 (16, 79000)	356 (2, 625)	602.5 (77, 1895)
Video content (n=100), median (IQR)						
Disease-related knowledge (n=87)	20 (27, 320000)	184 (1, 1927)	1081 (1, 29000)	3584 (10, 64000)	194.5 (2, 539)	273 (14, 3587)
Personal experience (n=11)	16064.5 (39, 64000)	1212 (0, 3308)	9393 (1, 44000)	25430 (35, 79000)	393 (2, 625)	550 (160, 1895)
Other (n=2)	5548.5 (4903, 16000)	1008 (520, 2536)	7803 (5394, 21000)	17000 (19000, 53000)	149.5 (136, 435)	3232.5 (1080, 7545)

Table 7. Distribution of GQS and five-level modified DISCERN reliability categories for TikTok and Bilibili videos.

Scale / category	TikTok (n=100), n (%)	Bilibili (n=100), n (%)
GQS		
1 (Very poor)	16 (16%)	8 (8%)
2 (Poor)	24 (24%)	32 (32%)
3 (Moderate)	21 (21%)	37 (37%)
4 (Good)	24 (24%)	19 (19%)
5 (Excellent)	15 (15%)	4 (4%)
Modified DISCERN reliability category		
1 (Unreliable; score 1)	4 (4%)	4 (4%)
2 (Less reliable; scores 2-3)	31 (31%)	26 (26%)
3 (Fairly reliable; scores 4-5)	30 (30%)	41 (41%)
4 (Relatively reliable; scores 6-7)	28 (28%)	26 (26%)
5 (Reliable; score 8)	7 (7%)	3 (3%)
Platform comparison: GQS	Mann-Whitney U	P=0.297
Platform comparison: modified DISCERN category	Mann-Whitney U	P=0.816

Table 8. Spearman correlations among engagement and temporal variables.

Variable	Likes	Comments	Saves	Shares	Days since published	Duration
Likes						
rho	1.0000	0.3444	0.6758	0.7443	0.1677	-0.0959
P value	<0.001	<0.001	<0.001	<0.001	0.0176	0.1767
Comments						
rho	0.3444	1.0000	0.7638	0.7429	0.6110	0.3965
P value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Saves						
rho	0.6758	0.7638	1.0000	0.9214	0.5190	0.2683
P value	<0.001	<0.001	<0.001	<0.001	<0.001	0.0001
Shares						
rho	0.7443	0.7429	0.9214	1.0000	0.4688	0.1686
P value	<0.001	<0.001	<0.001	<0.001	<0.001	0.0170
Days since published						
rho	0.1677	0.6110	0.5190	0.4688	1.0000	0.4271
P value	0.0176	<0.001	<0.001	<0.001	<0.001	<0.001
Duration						
rho	-0.0959	0.3965	0.2683	0.1686	0.4271	1.0000
P value	0.1767	<0.001	0.0001	0.0170	<0.001	<0.001

Table 9. Spearman correlations between video metrics and human-rated GQS and modified DISCERN reliability categories in the full 200-video dataset.

Video metric	GQS, rho	P value	Modified DISCERN category, rho	P value
Likes	-0.076	0.286	-0.081	0.253
Comments	-0.092	0.197	-0.038	0.594
Saves	-0.117	0.098	-0.076	0.288
Shares	-0.114	0.109	-0.072	0.310
Days since published	-0.108	0.127	-0.032	0.654
Duration	-0.008	0.905	-0.025	0.726
Followers	0.144	0.042	0.130	0.067

Table 10. Primary proportional-odds ordinal logistic regression for higher GQS category. OR>1 indicates higher odds of being in a higher GQS category.

Predictor	Adjusted OR for higher GQS	95% CI	P value	Reference / scale
Bilibili	0.83	0.43-1.60	0.582	TikTok
News agencies	1.27	0.60-2.71	0.534	Physicians
Independent users	0.29	0.13-0.63	0.002	Physicians
Video duration	1.24	0.92-1.67	0.156	Per SD of $\ln(1+x)$
Days since upload	0.81	0.56-1.17	0.270	Per SD of $\ln(1+x)$
Engagement index	0.93	0.68-1.28	0.663	Per SD

Table 11. Primary proportional-odds ordinal logistic regression for higher modified DISCERN reliability category. OR>1 indicates higher odds of being in a higher reliability category.

Predictor	Adjusted OR for higher modified DISCERN category	95% CI	P value	Reference / scale
Bilibili	1.03	0.53-1.99	0.941	TikTok
News agencies	1.31	0.59-2.90	0.513	Physicians
Independent users	0.40	0.18-0.88	0.023	Physicians
Video duration	1.10	0.80-1.50	0.569	Per SD of $\ln(1+x)$
Days since upload	0.89	0.61-1.29	0.531	Per SD of $\ln(1+x)$
Engagement index	0.95	0.69-1.31	0.757	Per SD

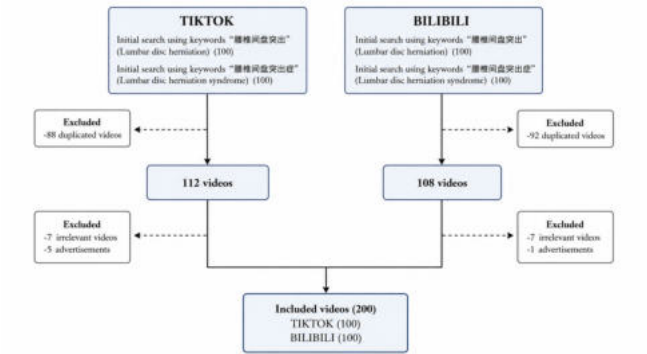


Figure 1. Search and screening process for lumbar disc herniation-related videos on TikTok and Bilibili.

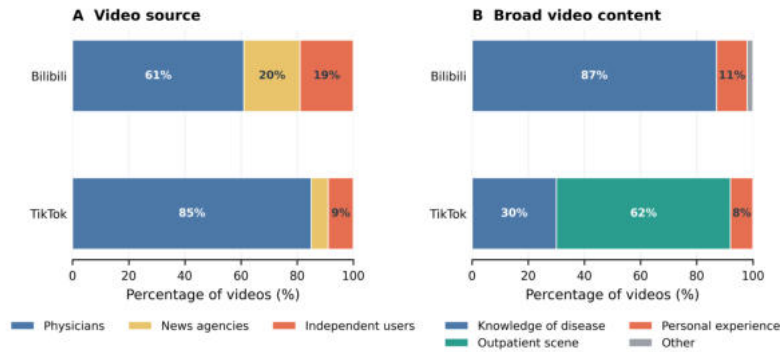


Figure 2. Platform composition of lumbar disc herniation-related videos. (A) Video source distribution. (B) Broad video content distribution. Percentages are based on 100 videos per platform.

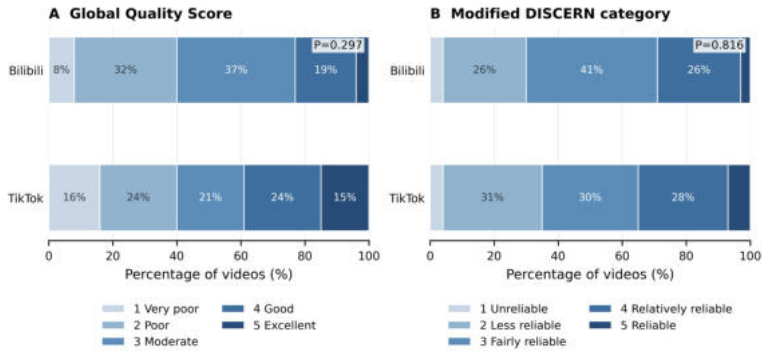


Figure 3. Ordinal distributions of information quality and reliability by platform. (A) Global Quality Score (GQS). (B) Five-level reliability category derived from the eight-item modified DISCERN score. P values are from two-sided Mann-Whitney U tests.

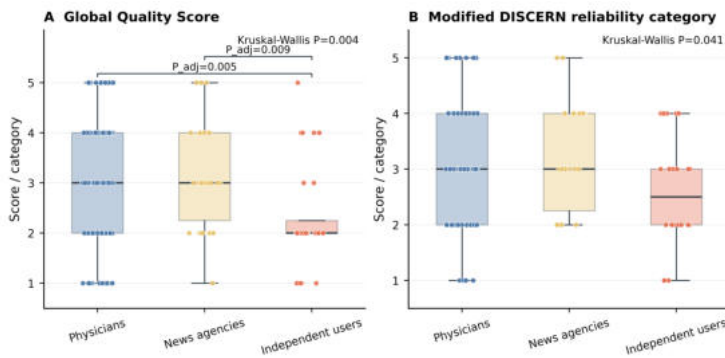


Figure 4. Human-rated GQS and modified DISCERN reliability categories according to video source. Boxes show the interquartile range with the median; points show individual videos. Overall P values are from Kruskal-Wallis tests. Pairwise GQS P values are Holm-adjusted Dunn comparisons.

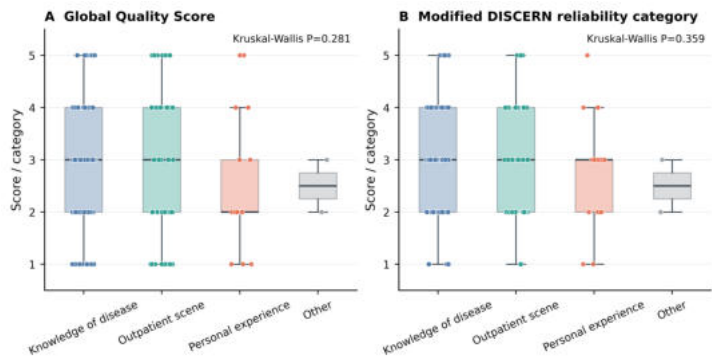


Figure 5. Human-rated GQS and modified DISCERN reliability categories according to broad video content. Boxes show the interquartile range with the median; points show individual videos. Overall P values are from Kruskal-Wallis tests.

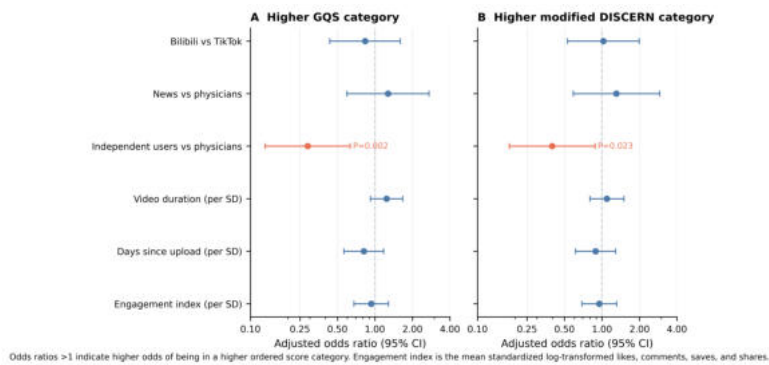


Figure 6. Multivariable proportional-odds ordinal logistic regression for higher (A) GQS and (B) modified DISCERN reliability categories. Adjusted odds ratios (ORs) are shown with 95% CIs. Primary models included platform, video source, log-transformed video duration, log-transformed days since upload, and a standardized engagement index.