

Assessing the quality and reliability of health information on GLP-1 receptor agonists on TikTok and Bilibili: a cross-sectional analysis of the public health infodemic

Gang Peng^{ID}, Chen Wang, Hong-Wei Zhang, Ting Xu and Jian-Zhong Di^{ID}

Ther Adv Endocrinol Metab

2026, Vol. 17: 1–16

DOI: 10.1177/
20420188261467887

© The Author(s), 2026.
Article reuse guidelines:
sagepub.com/journals-
permissions

Abstract

Background: Glucagon-like peptide-1 (GLP-1) receptor agonists have transitioned from clinical tools for diabetes to a global cultural phenomenon for weight loss, triggering an “infodemic.” While previous studies have characterized this landscape on Western platforms, the digital health ecosystem in China remains distinct. China is currently navigating a unique market shift driven by the June 2025 approval of the domestic dual agonist mazdutide, an impending 2026 patent cliff for semaglutide, and a specific market for counterfeit products.

Objectives: To assess and compare the quality, reliability, and content completeness of health information regarding GLP-1 agonists on China’s two dominant video platforms, Douyin (the Chinese version of TikTok) and Bilibili; to examine how platform architecture and regulatory context influence content reliability; and to contrast these findings with recent benchmarks from Western social media studies.

Design: Cross-sectional observational study.

Methods: A systematic search for “GLP1” was conducted on Douyin and Bilibili. Due to divergent algorithms, an asymmetric sampling strategy was used: all 106 videos returned by Douyin were screened, while the top 200 videos from Bilibili’s default ranking were assessed. After applying exclusion criteria, 128 videos (56 Douyin, 72 Bilibili) were analyzed. Quality was evaluated using the modified DISCERN tool, Global Quality Score, and JAMA benchmarks. No human subjects were involved; therefore, ethical approval was not required.

Results: Bilibili videos were significantly longer and demonstrated superior quality, scoring significantly higher than Douyin videos on both the DISCERN ($p < 0.001$) and JAMA benchmark ($p = 0.003$). Unlike findings on TikTok and Instagram, where quality was uniformly low, Bilibili hosted high-quality, in-depth content comparable to patient education materials. However, a safety analysis revealed videos promoting fraudulent products that co-opt the drug’s brand name, posing a severe risk unique to the unregulated domestic supply chain.

Conclusion: The Chinese digital landscape exhibits a quality tier on Bilibili that is absent in Western short-video counterparts. While Bilibili serves as a reliable “digital library,” Douyin remains a high-risk environment for misinformation and counterfeit sales. As China approaches the 2026 patent cliff, regulatory bodies must address the specific threat of fraudulent “GLP-1” branding on illicit products.

Trial registration: Not applicable.

Correspondence to:

Jian-Zhong Di
Department of Metabolic
and Bariatric Surgery,
Shanghai Sixth People’s
Hospital Affiliated to
Shanghai Jiao Tong
University School of
Medicine, Shanghai
200233, China
dijianzhong@sjtu.edu.cn

Gang Peng
Chen Wang
Hong-Wei Zhang
Ting Xu
Department of Metabolic
and Bariatric Surgery,
Shanghai Sixth People’s
Hospital Affiliated to
Shanghai Jiao Tong
University School of
Medicine, Shanghai, China

Plain language summary

Are social media videos about weight-loss drugs reliable? A comparison of information quality on TikTok and Bilibili in China

Weight-loss medications, such as semaglutide (Ozempic/Wegovy), tirzepatide (Mounjaro) and the newly approved Chinese drug mazdutide, have exploded in popularity. As demand grows, many people turn to social media video platforms like TikTok (Douyin) and Bilibili for medical advice. However, the accuracy of this information can vary greatly. This study analyzed 128 videos from these two major Chinese platforms to determine if they provide safe and reliable health information. The researchers found a significant difference between the two platforms. Bilibili, which hosts longer videos, acted like a “digital library,” offering detailed and accurate explanations of how these drugs work and their potential side effects. In contrast, TikTok videos were much shorter and often oversimplified complex medical details, making them less reliable for patients seeking deep understanding. Although many doctors use TikTok, the short video format limited the depth of their advice. Most concerning, the study found that a small number of videos on TikTok were promoting the sale of counterfeit or fake “GLP-1” products, such as “slimming teas” or unbranded injections. This poses a serious safety risk to consumers. The study suggests that while Bilibili is a better source for educational content, users on all platforms must be cautious of misinformation and illegal drug sales, especially as new drugs enter the Chinese market.

Keywords: Bilibili, counterfeit drugs, digital health, Douyin, GLP-1 receptor agonists, infodemic, mazdutide, semaglutide

Received: 19 December 2025; revised manuscript accepted: 12 June 2026.

Introduction

The GLP-1 receptor agonist phenomenon: From prescription drug to global sensation

Glucagon-like peptide-1 (GLP-1) receptor agonists are a class of medications that mimic the action of the endogenous incretin hormone GLP-1.¹ Their mechanism of action includes stimulating insulin secretion, suppressing glucagon release, delaying gastric emptying, and increasing satiety by acting on the central nervous system.² Initially developed for the management of type 2 diabetes, their significant efficacy in promoting weight loss has led to regulatory approvals for chronic weight management in adults and select adolescent populations.^{3–6}

In recent years, driven by celebrity endorsements and widespread social media coverage, GLP-1 agonists have transcended their clinical origins to become a global cultural phenomenon.⁷ This unprecedented attention has created a massive demand for information, with a growing number

of individuals turning to online platforms to learn about the efficacy, side effects, and accessibility of these drugs.⁸

Social media: The new frontier and the “infodemic”

The paradigm of health information acquisition has shifted from traditional sources, such as physicians and official health organizations, to digital platforms.^{9,10} Social media has become a prominent source of health information, particularly among younger demographics.^{11–13} A growing body of surveys further indicates that a substantial segment of the public now turns to social media influencers for health-related guidance.^{14,15}

The World Health Organization defines an “infodemic” as an overabundance of information—both accurate and inaccurate—that makes it difficult for people to find trustworthy sources.^{16–18} Recent research by Lee et al.¹⁹ characterized this infodemic on Western platforms—specifically

Instagram and TikTok International-finding uniformly low information quality dominated by non-medical influencers. However, the digital landscape in China presents a different ecosystem that remains under-researched.

The unique Chinese context: Mazdutide and the patent cliff

Critically, the content on Chinese platforms is shaped by specific regulatory and market forces distinct from those in the United States or Europe. A key driver is domestic innovation, highlighted by the National Medical Products Administration (NMPA)'s approval of mazdutide (Innovent Biologics, Suzhou, Jiangsu, China) in June 2025.²⁰ As the first domestic dual GLP-1/glucagon receptor agonist for chronic weight management, this "homegrown" competitor has sparked considerable attention in the medical community regarding its dual mechanism of action.²¹ Furthermore, the market is heavily influenced by the impending patent expiration for semaglutide in China, scheduled for 2026. This looming deadline has fueled significant speculation and, dangerously, a specific market for fraudulent products. Unlike the "compounding pharmacy" issues seen in the US, the Chinese market faces a distinct threat from items such as "GLP-1 herbal teas" or counterfeit devices that exploit the drug's brand name without containing any active ingredients.

Rationale and objectives

While previous studies like Lee *et al.*¹⁹ have evaluated health information quality on Western social media (specifically Instagram and TikTok International), a direct comparison of GLP-1 agonist content on platforms as structurally and culturally distinct as Douyin (the Chinese domestic TikTok) and Bilibili is currently lacking. This study aims to fill this gap by assessing the quality of information in the Chinese digital context, specifically highlighting how platform architecture, unique market factors, and regulatory context influence content reliability, and by contrasting these findings with those from Western social media platforms.

Terminology note

Throughout this paper, "TikTok" denotes the Chinese domestic version of the platform, Douyin, unless explicitly stated otherwise. By

contrast, the global version examined in previous Western studies is referred to as "TikTok International." This distinction is maintained to avoid conflation of the two separate platform ecosystems.

Methods

Search strategy and video selection

This cross-sectional study was designed to capture a snapshot of available content at a single point in time, and is reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for cross-sectional studies.²² We conducted a search using the keyword "GLP1" within the mobile applications of TikTok (i.e., Douyin, the Chinese domestic version) and Bilibili. To mitigate bias from personalized recommendation algorithms, all account information and search histories were cleared prior to the search.

The sampling methodology was asymmetric, reflecting the different search result mechanisms of the two platforms. On TikTok, the search yielded a finite total of 106 videos, all of which were collected for screening. On Bilibili, which returned a vast number of results without a displayed total count, we selected the first 200 videos returned by the platform's default search ranking. We note that Bilibili does not offer a "sort by popularity" or "most viewed" filter; the default ranking reflects the platform's standard algorithmic ordering, not a creator-curated quality sort. Random sampling was not feasible because Bilibili's infinite-scroll architecture does not provide an index of the full result set, and systematic scraping is prohibited by the platform's terms of service. This methodological difference implies that the Bilibili sample may be pre-filtered for popularity or quality by the platform's algorithm, whereas the TikTok sample represents the entirety of available content at that moment.

The exclusion criteria are detailed in Figure 1 and included: (1) irrelevant topic; (2) duplicate video; (3) no substantive health education content; (4) no sound; (5) posted within the last 7 days; and (6) published before 2020. Critically, videos found to be promoting or selling counterfeit GLP-1 products were also excluded, as this content constitutes a direct public health hazard.

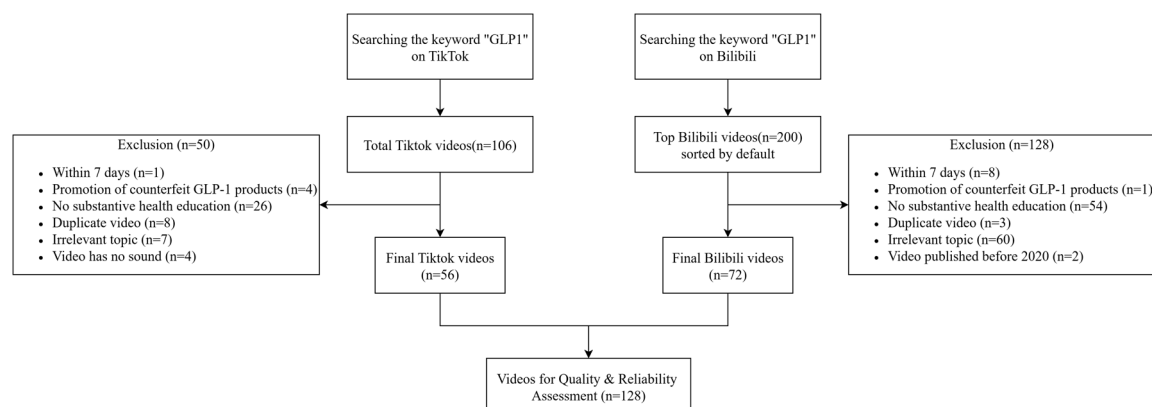


Figure 1. Flowchart of the video selection process. This diagram illustrates the methodology for video inclusion. It details the initial number of videos identified on TikTok ($n = 106$) and Bilibili ($n = 200$) using the keyword “GLP1.” It then specifies the number of videos excluded based on predefined criteria (e.g., irrelevant topic, promotion of counterfeit products, no substantive health education), leading to the final sample of 56 TikTok videos and 72 Bilibili videos for quality and reliability assessment. GLP1, glucagon-like peptide-1.

Data extraction

For each eligible video, metadata was extracted and compiled into a data file for analysis. The collected variables included: (1) Publisher Type: categorized as Personal User, Doctor (verified healthcare professional (HCP)), Organization (e.g., hospital, university, health department), or News Media based on account profiles and verification status; (2) Video Characteristics: year of publication, duration in seconds; and (3) User Engagement Metrics: number of likes, collections, comments, and retweets.

The extracted metadata, quality scores, and content completeness ratings for all 128 videos are provided in Table S1 (Raw Data Table).

Video quality and content assessment

Video quality was evaluated using three validated, standardized instruments, with detailed scoring criteria provided in the Supplemental Materials.

Modified DISCERN tool. A 5-item questionnaire assessing reliability, source transparency, and content balance. Each item is scored dichotomously (Yes = 1, No = 0), with total scores ranging from 0 to 5.^{23–25}

Global Quality Score. A 5-point Likert scale (1 = poor to 5 = excellent) providing an overall

assessment of video quality, flow, comprehensiveness, and utility for patients.^{26,27}

JAMA benchmarks. A 4-point checklist evaluating transparency and ethical standards, including disclosure of authorship, attribution, currency, and conflicts of interest. Each met criterion receives 1 point, for a total score of 0–4.²⁸

To assess content comprehensiveness, a bespoke “GLP-1 Content Completeness Assessment Form” was developed. This tool evaluates videos across six core educational dimensions: (1) Definition and Mechanism; (2) Indications; (3) Risk Factors and Contraindications; (4) Usage and Assessment; (5) Side Effects and Management; and (6) Outcomes and Efficacy. Each dimension was scored on a scale from 0 (no content) to 2 (extensive content), for a maximum total score of 12.

Two researchers with medical backgrounds independently scored all videos. To ensure consistency, inter-rater reliability was calculated for the initial, independent scores across all 128 videos. The agreement was found to be good. Cronbach’s alpha coefficients were 0.78 for the modified DISCERN tool, 0.83 for the Global Quality Score (GQS), 0.81 for the JAMA benchmarks, and 0.75 for the Content Completeness Assessment Form. Following this assessment, any scoring discrepancies were resolved by a

third, senior researcher to create the final dataset for analysis.

Statistical analysis

All statistical analyses were performed using R software (version 4.4.1).²⁹ Independent samples *t*-tests or Mann–Whitney *U* tests were used for two-group comparisons (e.g., TikTok vs Bilibili), while one-way ANOVA or Kruskal–Wallis tests were used for multi-group comparisons (e.g., publisher types). Spearman's rank correlation was used to assess the relationship between engagement metrics and quality scores. A two-tailed *p*-value of <0.05 was considered statistically significant.

Results

Video screening and sample characteristics

As illustrated in the flowchart (Figure 1), the initial search yielded 106 videos on TikTok and 200 on Bilibili. After applying exclusion criteria, 56 videos from TikTok and 72 from Bilibili were included in the final analysis, for a total of 128 videos.

Table 1 presents a comparison of the general characteristics of the videos from both platforms. Significant differences were observed in multiple domains. Bilibili videos were substantially longer, with 66.7% exceeding 300s, compared to only 16.1% on TikTok ($p < 0.001$). The distribution of publisher types also differed significantly ($p < 0.001$); doctors were the primary creators on TikTok (66.1%), whereas personal users were the main contributors on Bilibili (58.3%). Despite their shorter duration, TikTok videos garnered significantly higher median user engagement across likes, comments, and retweets ($p < 0.001$), reflecting the platform's high-intensity interaction model.

To illustrate the spectrum of content quality observed, Table 2 presents representative examples selected from the full dataset. These examples span high-quality, comprehensive content and low-quality, misleading, or potentially harmful content across both platforms. The complete raw data for all 128 videos is available in Table S1.

Content completeness analysis

The radar plots in Figure 2(a) and (b) illustrate that Bilibili videos scored higher on average across all six content dimensions (Definition and Mechanism, Indications, Risk Factors, Usage and Assessment, Side Effects and Management, and Outcomes and Efficacy), indicating more comprehensive coverage of the topic. Analysis of scores by year (Figure 2(c)–(n)) revealed no significant trend of improvement in content completeness on either platform from 2020 to 2025, suggesting that the quality of information has not kept pace with the topic's rising popularity.

Impact of publisher type on information quality

As shown in the heatmaps (Figure 3), videos created by doctors and organizations generally received higher mean scores across DISCERN, GQS, and JAMA criteria on both platforms, indicating superior quality from professional sources.

Table 3 provides a detailed breakdown by publisher type. While the differences in quality scores (DISCERN, GQS, JAMA) among publisher types did not reach statistical significance in this sample ($p > 0.05$), engagement metrics varied dramatically. Videos from doctors and news media attracted significantly more likes, comments, and retweets ($p < 0.001$). This suggests that while users may not accurately discern clinical quality, they tend to engage more with content from sources that appear authoritative.

Inter-platform comparison of video quality

The violin plots in Figure 4 display the central findings of this study. Bilibili videos achieved a significantly higher mean DISCERN score ($p < 0.001$) and a significantly higher mean JAMA benchmark score ($p = 0.003$) compared to TikTok. This indicates that content on Bilibili is generally more reliable, transparent, and better at disclosing authorship, sources, and potential conflicts of interest.

Interestingly, there was no statistically significant difference in GQS scores between the two platforms ($p = 0.148$). This suggests that while Bilibili excels in objective measures of reliability, users' subjective perceptions of overall quality and usefulness are comparable across both platforms.

Table 1. Overall characteristics of the videos.

Characteristic	Descriptive statistics		
	TikTok (<i>n</i> = 56)	Bilibili (<i>n</i> = 72)	<i>p</i>
Video duration (s, <i>n</i> (%))			<0.001
<60	12 (21.4)	1 (1.4)	
60–300	35 (62.5)	23 (31.9)	
>300	9 (16.1)	48 (66.7)	
Video upload year (<i>n</i> (%))			0.004
2020	0	1	
2021	0	1	
2022	3	4	
2023	2	15	
2024	8	20	
2025	43	31	
Video author type (<i>n</i> (%))			<0.001
Personal user	13	42	
Doctor	37	13	
Organization	0	3	
News media	6	14	
Are there any characters in the video? (<i>n</i> (%))			<0.001
Yes	49	38	
No	7	34	
Are there subtitles in the video? (<i>n</i> (%))			0.001
Yes	54	54	
No	2	18	
Is there background music in the video? (<i>n</i> (%))			0.059
Yes	21	16	
No	35	56	
Is there any animation in the video? (<i>n</i> (%))			0.254
Yes	31	47	
No	25	25	
Likes (<i>M</i> (<i>Q1</i> , <i>Q3</i>))	2283.00 (191.00, 11750.00)	25.50 (9.25, 137.50)	<0.001
Collections (<i>M</i> (<i>Q1</i> , <i>Q3</i>))	224.50 (23.50, 893.50)	33.50 (11.00, 141.75)	0.004
Comments (<i>M</i> (<i>Q1</i> , <i>Q3</i>))	959.00 (82.75, 5064.25)	6.00 (0, 79.25)	<0.001
Retweets (<i>M</i> (<i>Q1</i> , <i>Q3</i>))	1049.00 (74.75, 7381.50)	17.00 (3.00, 76.25)	<0.001

Table 2. Representative examples of analyzed videos.

Video ID	Platform	Publisher type	Title (abbreviated)	DISCERN	GQS	JAMA
7	TikTok	Doctor	GLP-1 three giants: semaglutide vs tirzepatide vs mazdutide	5	5	4
97	Bilibili	Doctor	Comprehensive analysis of the weight-loss drug semaglutide	5	5	4
1	TikTok	Personal User	Discovered a new once-weekly GLP-1	1	2	1
89	Bilibili	Personal user	New weight-loss drug 1.5× more effective than semaglutide?	2	2	2

GLP-1, glucagon-like peptide-1; GQS, Global Quality Score.

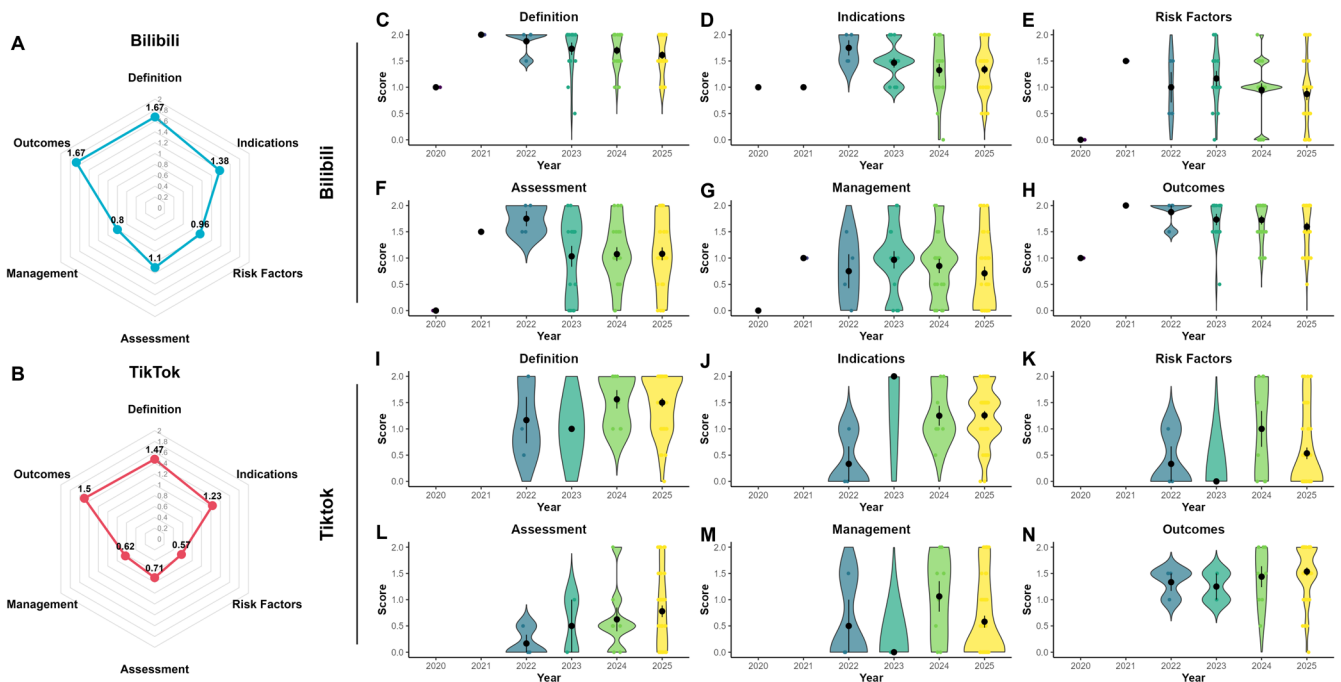


Figure 2. Video content completeness assessment. This figure evaluates the comprehensiveness of video content. (a, b) are radar plots showing the mean scores of videos on Bilibili and TikTok, respectively, across six key educational dimensions: Definition, Indications, Risk Factors, Assessment, Management, and Outcomes. (c–n) are violin plots that display the distribution of scores for each of these six dimensions, broken down by the year of publication (2020–2025), allowing for a temporal analysis of content quality on both platforms.

Correlation between user engagement and video quality

The correlation matrices in Figure 5 reveal distinct patterns on each platform. On Bilibili (Figure 5(a)), video duration was the only variable that showed a significant positive correlation with all three quality scores (DISCERN, GQS, and JAMA). Engagement metrics such as likes

and comments showed no significant correlation with quality. This suggests that on Bilibili, content depth (proxied by duration) is a reliable indicator of quality, while popularity is not.

In stark contrast, on TikTok (Figure 5(b)), likes, collections, and retweets, in addition to duration, were all significantly and positively correlated

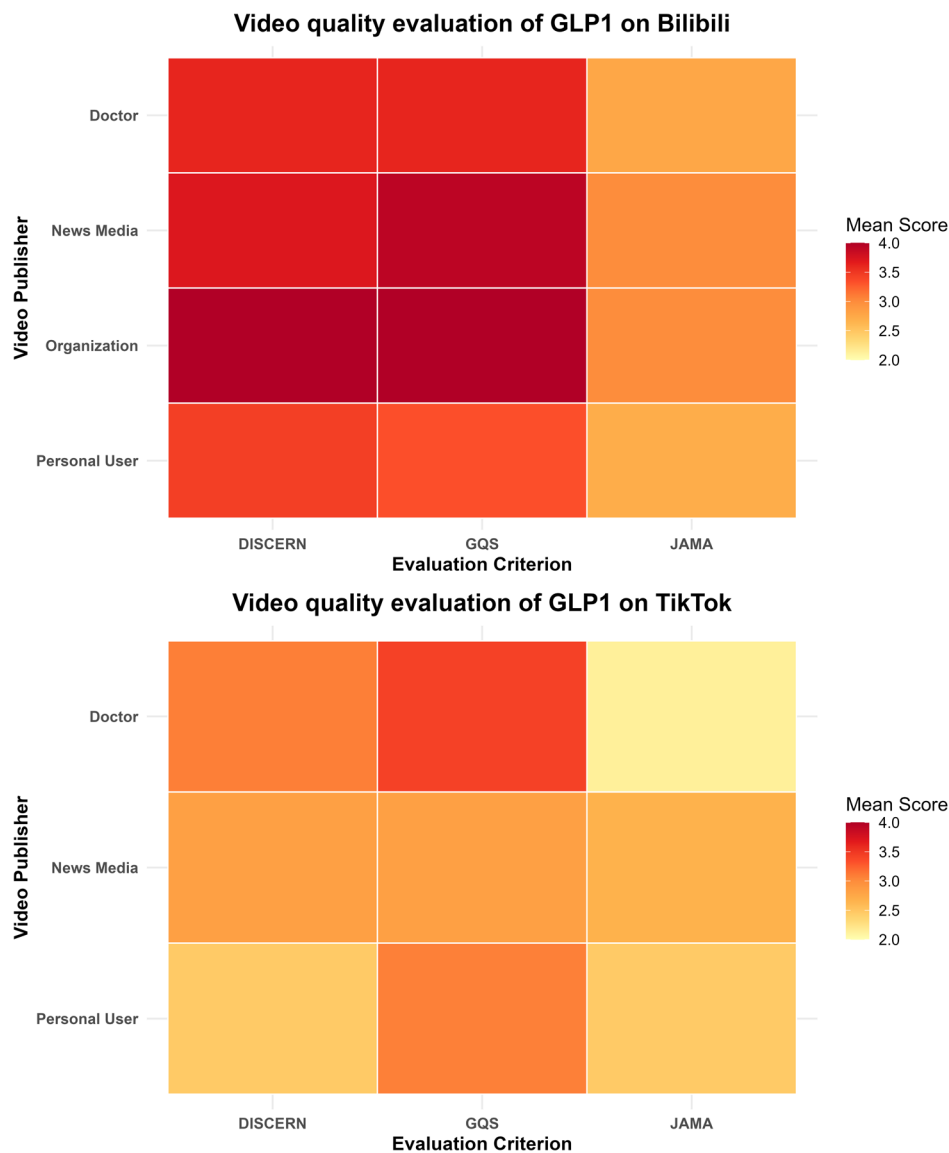


Figure 3. Heatmap of video quality evaluation by publisher type. This figure visualizes the average quality scores for videos based on the type of creator. The top panel shows data for Bilibili and the bottom panel for TikTok. Publishers are categorized as Doctor, News Media, Organization, and Personal User. The scores for DISCERN, GQS, and JAMA benchmarks are represented by color intensity, where darker red indicates a higher mean score and thus better video quality. GQS, Global Quality Score.

with DISCERN and GQS scores. This indicates that on TikTok, higher-quality content is more likely to receive higher engagement, although the reverse is not necessarily true.

Table 4 reinforces the importance of video length. When videos were categorized by duration, those longer than 300s scored significantly higher on all three quality metrics—DISCERN, GQS, and JAMA—than shorter videos (all $p < 0.001$). This

provides strong evidence that adequate time is necessary to convey complex medical information accurately and comprehensively.

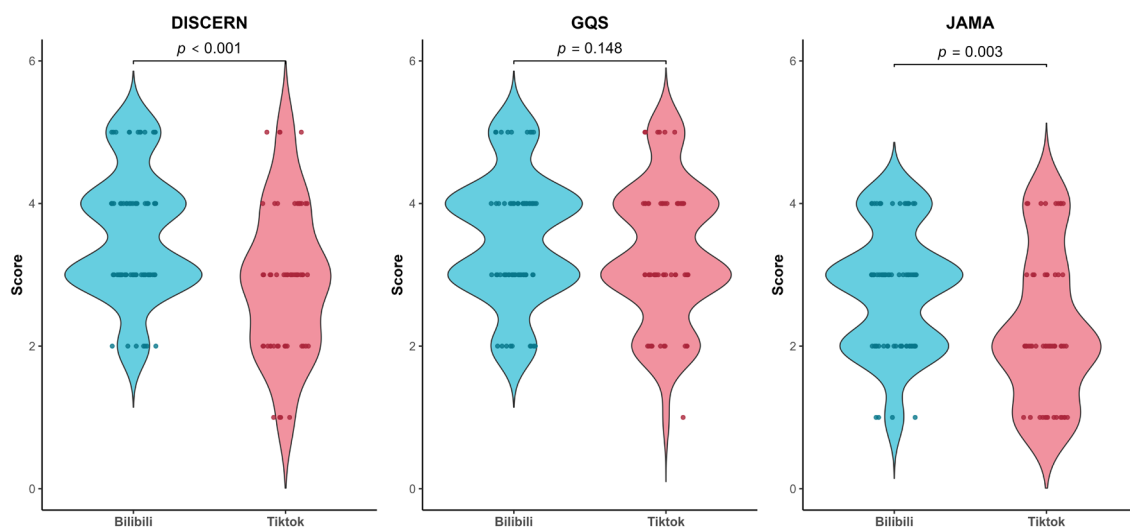
Prevalence of harmful content:
Promotion of counterfeit drugs

As noted in the screening process (Figure 1), a total of five videos were excluded because their content involved the promotion or sale of

Table 3. Characteristics of GLP1 videos by video publisher type.

Variables	Total (n = 128)	Personal user (n = 55)	Doctor (n = 50)	News media (n = 20)	Organization (n = 3)	p
Likes, <i>M</i> (Q1, Q3)	148.00 (17.75, 3719.50)	25.00 (10.00, 994.00)	622.50 (130.75, 6145.75)	168.00 (14.75, 12687.00)	30.00 (16.50, 33.50)	<0.001
Collections, <i>M</i> (Q1, Q3)	55.00 (11.75, 388.50)	33.00 (8.50, 177.50)	116.00 (34.25, 540.75)	89.50 (8.75, 3483.75)	31.00 (21.00, 40.50)	0.049
Comments, <i>M</i> (Q1, Q3)	58.50 (3.00, 1217.50)	5.00 (1.00, 154.50)	219.50 (44.25, 3453.50)	103.00 (2.25, 2465.50)	3.00 (2.00, 9.00)	<0.001
Retweets, <i>M</i> (Q1, Q3)	67.00 (8.75, 1484.50)	21.00 (3.00, 221.00)	363.50 (46.25, 3593.25)	97.00 (9.50, 7119.50)	18.00 (12.50, 19.00)	<0.001
Duration (s), <i>M</i> (Q1, Q3)	252.50 (95.50, 674.50)	434.00 (210.50, 1089.50)	124.50 (78.00, 255.00)	348.50 (152.50, 891.75)	579.00 (331.50, 614.50)	<0.001
DISCERN, mean (SD)	3.27 (1.00)	3.22 (1.03)	3.22 (0.91)	3.45 (1.05)	4.00 (1.73)	0.522
GQS, mean (SD)	3.42 (0.93)	3.27 (0.89)	3.48 (0.81)	3.60 (1.23)	4.00 (1.00)	0.283
JAMA, mean (SD)	2.56 (0.97)	2.65 (1.04)	2.30 (0.89)	2.90 (0.85)	3.00 (1.00)	0.063

GLP-1, glucagon-like peptide-1; GQS, Global Quality Score.

**Figure 4.** Comparison of video quality scores between Bilibili and TikTok. This figure presents violin plots comparing the distribution and mean scores of videos from Bilibili and TikTok on three validated quality assessment tools: DISCERN, GQS, and JAMA benchmarks. *p*-Values are displayed to indicate the statistical significance of the differences in quality scores between the two platforms.

GQS, Global Quality Score.

counterfeit GLP-1 products. Four of these videos were found on TikTok (representing 3.8% of its initial sample), and one was found on

Bilibili (0.5% of its initial sample). This finding highlights a direct and serious public safety risk.

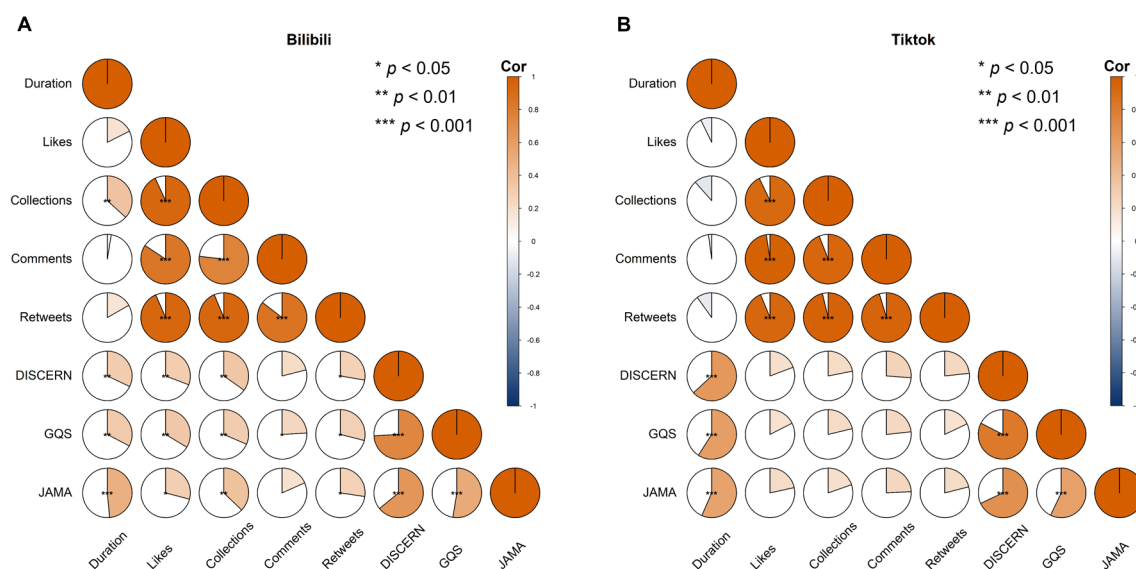


Figure 5. Correlation matrix of video characteristics and quality scores. This figure displays two correlograms showing the Spearman correlation coefficients between various video metrics (Duration, Likes, Collections, Comments, Retweets) and the three quality scores (DISCERN, GQS, JAMA). Panel (a) represents Bilibili, and panel (b) represents TikTok. The strength and direction of correlations are indicated by the size and color of the circles (darker orange signifies a stronger positive correlation). Asterisks denote the level of statistical significance.

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

GQS, Global Quality Score.

Table 4. Characteristics of GLP1 videos by duration category.

Variables	Total ($n = 128$)	Duration < 60 s ($n = 13$)	Duration 60–300 s ($n = 58$)	Duration > 300 s ($n = 57$)	p
Likes, M (Q1, Q3)	148.00 (17.75, 3719.50)	1743.00 (200.00, 15,000.00)	138.50 (25.25, 3392.00)	79.00 (13.00, 1765.00)	0.041
Collections, M (Q1, Q3)	55.00 (11.75, 388.50)	311.00 (15.00, 1924.00)	35.50 (7.75, 246.75)	80.00 (22.00, 488.00)	0.066
Comments, M (Q1, Q3)	58.50 (3.00, 1217.50)	307.00 (58.00, 6477.00)	84.00 (11.00, 1594.75)	18.00 (1.00, 472.00)	0.006
Retweets, M (Q1, Q3)	67.00 (8.75, 1484.50)	1058.00 (102.00, 14,000.00)	62.50 (15.50, 1383.50)	36.00 (7.00, 329.00)	0.014
DISCERN, mean (SD)	3.27 (1.00)	1.92 (0.64)	3.05 (0.83)	3.81 (0.85)	< 0.001
GQS, mean (SD)	3.42 (0.93)	2.23 (0.73)	3.26 (0.64)	3.86 (0.93)	< 0.001
JAMA, mean (SD)	2.56 (0.97)	1.46 (0.66)	2.19 (0.78)	3.19 (0.77)	< 0.001
GLP-1, glucagon-like peptide-1; GQS, Global Quality Score.					

Discussion

Principal findings

This study provides a systematic comparison of health information quality regarding GLP-1 receptor agonists on TikTok and Bilibili. The key findings are fourfold. First, a significant quality gap exists between the platforms, with Bilibili offering more reliable and comprehensive content than TikTok. This sharply contrasts with recent findings by Lee *et al.*¹⁹ on Western platforms (TikTok International and Instagram), where information quality was found to be uniformly poor across the board. Second, the creator's identity is a critical determinant of quality, with Chinese HCPs maintaining a higher standard of evidence-based communication than the “influencer-dominated” landscape observed in the West. Third, user engagement metrics are not reliable proxies for information quality, particularly on Bilibili. Finally, and most alarmingly, both platforms host dangerous content promoting fraudulent products that co-opt the “GLP-1” brand name, posing a direct threat to public health.

Platform architecture as a determinant of content quality

Our findings support a “digital environment” hypothesis, wherein a platform's intrinsic architecture, community culture, and business model profoundly shape its health content ecosystem.³⁰ Bilibili's superior performance can be attributed to its platform characteristics. It natively supports longer-form videos (Table 1), and greater duration is a prerequisite for discussing complex pharmacological mechanisms, such as the dual-agonist action of the newly approved mazdutide. Furthermore, Bilibili's user base consists predominantly of young, digitally native demographics with a strong and documented preference for in-depth, knowledge-oriented content, creating a fertile ground for high-quality information.^{31,32} The correlation analysis (Figure 5) confirms this: on Bilibili, only duration—a proxy for depth—correlates with quality, while superficial engagement metrics are decoupled from it.

Conversely, TikTok's architecture prioritizes short, fast-paced, and highly engaging content, with an algorithm designed to maximize virality.³³ This environment is inherently less suited for the nuanced and detailed explanations required for medical education. Although doctors were the

most frequent creators on TikTok in our sample (Table 1), the short-video format constrains the depth and completeness of the information they can convey, resulting in lower scores on objective quality assessments. This aligns with the limitations identified by Lee *et al.*¹⁹ (who studied TikTok International and Instagram), suggesting that the short-video format itself acts as a barrier to quality, regardless of the creator's credentials.

To better understand why health information quality diverges so significantly between the two platforms, Table 5 synthesizes our empirical findings with established platform architectures, algorithmic logics, and demographic characteristics documented in recent literature.

These platform-level differences must also be understood within the broader regulatory environment. In the United States, Section 230 of the Communications Decency Act shields platforms from liability for user-generated content, creating no legal obligation to verify the credentials of health influencers.⁴² This likely explains why Lee *et al.*¹⁹ found uniformly poor information quality on Instagram and TikTok International. In contrast, China introduced two landmark policies in late 2025: a joint multi-agency notice requiring platforms to authenticate medical licenses of health-content creators, and a regulation obliging influencers to hold professional qualifications before discussing medical topics.^{43–45} Such “gate-keeping” mandates filter out unqualified sources at the account level. This institutional mechanism explains why physicians constituted 66.1% of our TikTok sample, in sharp contrast to the Western influencer-dominated landscape. The European Union's Digital Services Act occupies a middle ground, demanding systemic risk mitigation but not pre-publication vetting.⁴⁶ The persistence of counterfeit promotions in our sample (3.8% on TikTok) reflects the fact that these policies were newly enacted and enforcement is still in its transitional phase, a pattern consistent with “fragmented authoritarianism” in Chinese health internet governance.⁴⁷

Sanctions for factual misrepresentation further amplify these regulatory differences. In the United States, Section 230 grants platforms broad immunity; no federal statute imposes direct fines on platforms or creators for spreading health misinformation. The EU's Digital Services Act can fine very large platforms up to 6% of their annual

Table 5. Platform characteristics of Bilibili and TikTok—a literature-based comparison.

Characteristics	TikTok (Douyin)	Bilibili	Source
Primary content focus	Short-form entertainment, trending topics, memetic imitation	Anime, gaming, knowledge-sharing, popular science videos	[31, 34]
Typical video duration	Predominantly short-form (83.9% <5 min)	Medium-to-long form (66.7% >5 min)	Current study
Signature feature	Algorithm-driven FYP	Danmaku—real-time scrolling overlaid comments	[35, 36]
Target demographics	Highly adopted across mass market, especially among adolescents	Predominantly Gen Z; e.g., 83% of users aged 18–24	[32, 37]
User engagement model	High volume of likes, comments, and fast-paced viral spread	Interactive learning, “Sanlian” (like, coin, collect), social presence	Current study, [38]
Algorithm and distribution	Virality, instant stimulation, algorithmized self-making	Community-driven, subcultural territory, creator (UP host) stickiness	[35, 39]
E-commerce integration	Highly integrated live-streaming commerce and product exposure	Moderate, primarily dependent on content creator sponsorships	[40]
Content creator types	Influencers, verified doctors, news media (in current sample)	Personal users, hobbyists, professional “UP hosts”	Current study, [41]
Community culture	Attention economy, trending hashtags, visibility labors	Strong subculture identity, carnivalesque, fan cohesion	[39, 41]
FYP, “For You” page.			

global turnover for systemic failures to mitigate risks, but it does not mandate pre-publication vetting.⁴⁸ China’s 2025 regulations, by contrast, impose direct monetary penalties on individual creators (up to 100,000 yuan) and on platforms that fail to verify credentials or remove illegal content, with additional business-suspension powers.⁴⁵ This stricter sanction regime reinforces the gatekeeping effect, making the production and dissemination of unverified health claims both legally risky and economically unattractive, which likely contributes to the higher baseline quality observed on Chinese platforms.

The “credibility gap” and the role of HCPs

This study highlights a “credibility gap” in the social media health landscape. While high-quality content from HCPs exists (Figure 3), it is often diluted by a larger volume of low-quality, anecdotal, or misleading content from personal users. On Bilibili, despite its higher overall quality, personal users were the dominant content creators

(Table 1), increasing the challenge for users to identify reliable information.

Research shows that content from HCPs is generally of higher quality but often struggles to compete for views and engagement against more sensational or entertaining videos from non-medical influencers.⁴⁹ Our data reflects this tension: while content from doctors is of higher quality, the immense user engagement on TikTok is driven more by the easily consumable short-form format than by the accuracy of the information presented.⁵⁰

A grave public health threat: The online market for fraudulent “GLP-1” products

The most alarming finding of this study is the presence of videos directly promoting and facilitating the sale of counterfeit and fraudulent products on mainstream social media platforms. This issue is distinct from the “compounding pharmacy” debate in the United States; in the Chinese

context, it is driven by the speculative atmosphere surrounding the impending 2026 patent expiration for semaglutide.

Unscrupulous sellers are exploiting the public's anticipation of "generic" options by marketing completely fraudulent items—such as "GLP-1 herbal teas," "slimming enzymes," or unbranded liquids—that simply appropriate the drug's name without containing any active pharmaceutical ingredient. This elevates the health information problem from "misinformation" (which may cause indirect harm) to the "enablement of illicit and dangerous transactions" (which can cause direct, severe harm).^{51,52} Regulatory bodies like the NMPA have recently cracked down on illicit sales, yet our analysis found that 3.8% of the initial search results on TikTok fell into this dangerous category. This means an average user seeking legitimate medical information has a non-trivial chance of encountering content that could lead them to purchase ineffective or potentially toxic fakes. This poses a formidable challenge not only to individual safety but also to global drug regulation and public health.⁵³

AI-generated content: An emerging regulatory frontier

A further dimension that this study could not systematically address is the role of AI-generated content (AIGC). At the time of data collection, China's mandatory AI-labeling regulation—the Administrative Provisions on the Labeling of AI-Generated and Synthetic Content, effective September 1, 2025—had only recently come into force, requiring platforms such as Douyin and Bilibili to ensure AIGC carries both visible markers and invisible digital watermarks.^{54–56} Consequently, our sample lacks explicit AI labels that would have enabled systematic analysis. Of note, recent studies have documented that AI-generated deepfakes are already being used on Western platforms to promote fraudulent GLP-1 products, underscoring the urgency of addressing AIGC in future research.^{57,58} Our findings provide a pre-regulation baseline against which the impact of China's new labeling mandate can be measured.

Study limitations

This study has several limitations. First, the cross-sectional design captures data at a single point in time and cannot account for the dynamic

evolution of social media content and algorithms. Second, the selection of the first 200 videos from Bilibili's default search ranking may introduce selection bias, though the direction of this bias is uncertain—the default ranking is not explicitly a popularity or quality sort, and the platform's algorithmic logic is not publicly documented. Third, the study focused exclusively on Chinese-language videos, which may limit the generalizability of the findings to other linguistic and cultural contexts. Fourth, our assessment was limited to the video content itself and did not investigate viewer interpretation, comprehension, or subsequent impact on health behaviors. Finally, the sampling window precluded analysis of AI-labeled content, as mandatory AIGC labeling requirements had only recently been implemented and were not yet consistently enforced.

Conclusion

This comparative analysis demonstrates that Bilibili is a relatively more reliable source for information on GLP-1 receptor agonists than TikTok, largely due to its support for longer, more detailed content. However, neither platform consistently provides high-quality medical information. Based on these findings, we call for coordinated action. Platform operators must implement stricter content moderation policies, specifically targeting the sale of illegal pharmaceuticals and fraudulent "GLP-1" branded products, while collaborating with public health agencies to algorithmically amplify verified information. HCPs should actively engage on these platforms by creating accessible content tailored to each format—such as in-depth explanations on Bilibili and myth-busting shorts on TikTok—to proactively counter misinformation. Simultaneously, public health organizations should launch targeted digital health literacy campaigns to educate the public on evaluating online health claims, expanding upon models of collaboration with social media entities. Finally, future research must evaluate the effectiveness of these interventions and conduct longitudinal studies to track the evolution of this health infodemic.

Declarations

Ethics approval and consent to participate

This study utilized exclusively publicly available data from social media platforms (TikTok and

Bilibili). It did not involve human participants, interventions, or the collection of private identifiable data. Therefore, in accordance with standard ethical guidelines for internet-mediated research, the observation of public online behavior where there is no reasonable expectation of privacy does not constitute human subjects research. Consequently, ethical approval from an Institutional Review Board (IRB) was not required. Consent to participate: Not applicable.

Consent for publication

Not applicable.

Author contributions

Gang Peng: Conceptualization; Data curation; Formal analysis; Methodology; Visualization; Writing – original draft.

Chen Wang: Data curation; Formal analysis; Methodology; Writing – review & editing.

Hong-Wei Zhang: Data curation; Formal analysis; Methodology; Writing – review & editing.

Ting Xu: Investigation; Methodology; Writing – review & editing.

Jian-Zhong Di: Funding acquisition; Project administration; Supervision; Writing – review & editing.

Acknowledgments

None.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Noncommunicable Chronic Diseases-National Science and Technology Major Project (Grant No. 2025ZD0550504), the National Natural Science Foundation of China (Grant No. 82370901), and the Science and Technology Commission of Shanghai Municipality (Grant No. 24DZ2304800).

Competing interests

The authors declare that there is no conflict of interest.

Availability of data and materials

The dataset supporting the conclusions of this article is available from the corresponding author on reasonable request.*

ORCID iDs

Gang Peng  <https://orcid.org/0009-0003-5968-3678>

Jian-Zhong Di  <https://orcid.org/0000-0003-4932-7189>

Supplemental material

Supplemental material for this article is available online.

References

1. Müller TD, Finan B, Bloom SR, et al. Glucagon-like peptide 1 (GLP-1). *Mol Metab* 2019; 30: 72–130.
2. Park JS, Kim KS and Choi HJ. Glucagon-like peptide-1 and hypothalamic regulation of satiation: cognitive and neural insights from human and animal studies. *Diabetes Metab J* 2025; 49: 333–347.
3. Nauck MA. Glucagon-like peptide 1 (GLP-1) in the treatment of diabetes. *Horm Metab Res* 2004; 36: 852–858.
4. Tran K, Park YI, Pandya S, et al. Overview of glucagon-like peptide-1 receptor agonists for the treatment of patients with type 2 diabetes. *Am Health Drug Benefits* 2017; 10(4): 178–188.
5. Zheng Z, Zong Y, Ma Y, et al. Glucagon-like peptide-1 receptor: mechanisms and advances in therapy. *Signal Transduct Target Ther* 2024; 9: 234.
6. Moiz A, Filion KB, Tsoukas MA, et al. Mechanisms of GLP-1 receptor agonist-induced weight loss: a review of central and peripheral pathways in appetite and energy regulation. *Am J Med* 2025; 138: 934–940.
7. Yi D. The global shortage of glucagon-like peptide-1 receptor agonists. *J Korean Diab* 2024; 25: 1–3.
8. Fong S, Carollo A, Lazuras L, et al. Ozempic (glucagon-like peptide 1 receptor agonist) in social media posts: unveiling user perspectives through Reddit topic modeling. *Emerg Trends Drugs Addict Health* 2024; 4: 100157.
9. Erlandsson LC, Guijo MS and Quintana-Alonso R. Decoding patterns: the crucial role of social media in health information consumption and user dynamics among the general population. *J Public Health* 2024; 34(5): 1–7.
10. Zongmin J, Huilin LIU, Wei JIN, et al. Status of health communication of medical institutions

- using new media platforms in Shanghai. *Shanghai J Prev Med* 2024; 36(03): 269–273.
11. Rowdhwal SSS and El Boghdady M. 150 The influence of TikTok on enhancing surgical patient education: a systematic review. *Br J Surg* 2025; 112(Suppl_10): znaf128.653).
 12. Auxier B and Anderson M. *Social media use in 2021*. Washington, DC: Pew Research Center, 2021.
 13. Swire-Thompson B and Lazer D. Public health and online misinformation: challenges and recommendations. *Annu Rev Public Health* 2020; 41: 433–451.
 14. Obutuzi NG. Social media and health trends: analyzing public perception. *Newport Int J Res Med Sci* 2025; 6: 163–170.
 15. Wang Y, McKee M, Torbica A, et al. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med* 2019; 240: 112552.
 16. Sivakumar A and Boratne AV. Infodemic: a blessing or curse? *Int J Community Med Public Health* 2022; 9: 1563.
 17. Patil S, Gondhali G, Choudhari S, et al. Infodemic—a new rapidly evolving virtual communicable pandemic with global threat! Hypothetical or real? *World J Adv Pharm Med Res* 2023; 4: 12–31.
 18. Gentili A, Villani L, Osti T, et al. Strategies and bottlenecks to tackle infodemic in public health: a scoping review. *Front Public Health* 2024; 12: 1438981.
 19. Lee S, Narula N, Weglarz M, et al. Evaluation of social media videos on weight-loss surgery and GLP-1 agonists: a content quality and reliability analysis. *Surg Endosc* 2025; 39: 6032–6036.
 20. Shirley M. Mazdutide: first approval. *Drugs* 2025; 85: 1621–1627.
 21. Ji L, Jiang H, Cheng Z, et al. A phase 2 randomised controlled trial of mazdutide in Chinese overweight adults or adults with obesity. *Nat Commun* 2023; 14: 8289.
 22. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007; 370: 1453–1457.
 23. Charnock D, Shepperd S, Needham G, et al. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999; 53: 105–111.
 24. Shan Y, Xing Z, Dong Z, et al. Translating and adapting the DISCERN instrument into a simplified Chinese version and validating its reliability: development and usability study. *J Med Internet Res* 2023; 25: e40733.
 25. Allam A, Schulz PJ and Krauthammer M. Toward automated assessment of health Web page quality using the DISCERN instrument. *J Am Med Inform Assoc* 2017; 24: 481–487.
 26. Bernard A, Langille M, Hughes S, et al. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. *Am J Gastroenterol* 2007; 102: 2070–2077.
 27. Yeung AWK. Evaluation of content quality of online health information by global quality score: a case study of researchers misnaming it and citing secondary sources. *Publications* 2025; 13: 23.
 28. Silberg WM, Lundberg GD and Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveat lector et viewer—let the reader and viewer beware. *JAMA* 1997; 277: 1244–1245.
 29. R Core Team R. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2020.
 30. Ruokolainen J, Nätti S, Juutinen M, et al. Digital healthcare platform ecosystem design: a case study of an ecosystem for Parkinson's disease patients. *Technovation* 2023; 120: 102551.
 31. Tianyi M. Exploring online learning with educational science videos: a case study of an educational video blogger on Bilibili. In: *Proceedings of the 2023 3rd international conference on social development and media communication (SDMC 2023)*, 30 December 2023, pp. 231–237. Dordrecht, The Netherlands: Atlantis Press.
 32. Qi X, Lu Q, Li B, et al. Videos in short video sharing platforms as a source of information on bipolar disorder: a cross-sectional content analysis study. *Front Public Health* 2025; 13: 1627885.
 33. Mulcahy R, Barnes R, de Villiers Scheepers R, et al. Going viral: sharing of misinformation by social media influencers. *Australas Mark J* 2024; 33: 296–309.
 34. Zulli D and Zulli D. Extending the Internet meme: conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media Soc* 2022; 24: 1872–1890.

35. Bhandari A and Bimo S. Why's everyone on TikTok now? The algorithmized self and the future of self-making on social media. *Soc Media + Society* 2022; 8: 205630512210862.
36. Chen Y, Gao Q and Rau P-L. Understanding gratifications of watching Danmaku videos—videos with overlaid comments. In: cross-cultural design methods, practice and impact: 7th international conference, CCD 2015, Held as Part of HCI International 2015, 2015, pp. 153–163. Springer, Cham.
37. Montag C, Yang H and Elhai JD. On the psychology of TikTok use: a first glimpse from empirical findings. *Front Public Health* 2021; 9: 641673.
38. Lu H, Limniou M and Zhang X. Metacognition and social presence in connectivist learning: an analysis of Bilibili interactions. *Educ Sci* 2025; 15: 1673.
39. Yin Y and Fung A. Youth online cultural participation and Bilibili. In: Luppigini R and Baarda R (eds) Digital media integration for participatory democracy. Hershey, PA: IGI Global; 2017, pp. 130–154.
40. Wang Y, Lu Z, Cao P, et al. How live streaming changes shopping decisions in E-commerce: a study of live streaming commerce. *Comput Supported Coop Work* 2022; 31: 701–729.
41. Abidin C. Mapping internet celebrity on TikTok: exploring attention economies and visibility labours. *Cult Sci J* 2021; 12: 77–103.
42. Communications Decency Act of 1996, 47 U.S.C. § 230 (1996).
43. Wang X. China set to clean up online health content. *People's Daily*, 4 August 2025.
44. Wang X. China steps up supervision over online health science popularization. *China Daily*, 2 August 2025.
45. Danao M. China's new influencer crackdown bans unqualified voices from covering these “sensitive topics” to fight misinformation. *Moneywise*, 4 November 2025.
46. European Union. Regulation (EU) 2022/2065 of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act). *Official J Law* 2022; 277: 1–102.
47. Luo T and Lv A. “Nine dragons run the water”: Fragmented internet governance in China. In: Haggart B, Tusikov N and Scholte JA (eds) *Power and authority in internet governance: return of the state?* London, UK: Routledge, 2021, pp. 123–146.
48. Busch C. *Platform responsibility in the European Union*. Cambridge, UK: Cambridge University Press, 2025.
49. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. *P T* 2014; 39: 491–520.
50. Etta RE, Babatunde AO, Okunlola PO, et al. The assessment of TikTok as a source of quality health information on human papillomavirus: a content analysis. *Cureus* 2024; 16: e75419.
51. Ashraf AR, Mackey TK, Schmidt J, et al. Safety and risk assessment of no-prescription online semaglutide purchases. *JAMA Netw Open* 2024; 7: e2428280.
52. Turnock LA, Hearne E and Lazuras L. Made in China: the international supply of illicit semaglutide and weight-loss medicines online. *Emerg Trends Drugs Addict Health* 2025; 5: 100169.
53. Ofori-Parku SS. Fighting the global counterfeit medicines challenge: a consumer-facing communication strategy in the US is an imperative. *J Glob Health* 2022; 12: 03018.
54. Cyberspace Administration of China, Ministry of Industry and Information Technology, Ministry of Public Security, National Radio and Television Administration. Administrative provisions on the labeling of AI-generated and synthetic content, (2025).
55. Douyin. Douyin announcement on upgrading AI content labeling function, <https://www.douyin.com/> (2025, accessed 4 May 2026).
56. Bilibili. Community rules update: AI-generated content identification function, <https://www.bilibili.com/> (2025, accessed 4 May 2026).
57. Fake AI-generated ads are on the rise, and TikTok is helping to create them. *Irish Examiner*, 14 December 2025.
58. Propfe LE and Seifert R. Misrepresentation of semaglutide in social media. *Naumyn Schmiedebergs Arch Pharmacol* 2026; 399: 815–832.