Jurnal Komunikasi: Malaysian Journal of Communication
https://ejournal.ukm.my/mjc
Volume 41(4) 2025: 230-246

Quality vs. Reach in Health Short Videos: A Dual-Path Test of the Heuristic–Systematic Model

SHAN CHEN
XI XI ZHAO
EMMA MIRZA WATI MOHAMAD*
ARINA ANIS AZLAN
Universiti Kebangsaan Malaysia

ABSTRACT

Short-video platforms such as Douyin broaden access to health information but may reward visibility over verifiability. Guided by the heuristic–systematic model, this study asks whether the same content features that improve information quality also help videos reach larger audiences. We retrieved 300 diabetes-related videos from Douyin's "For You" feed in three logged-out collection waves, and after removing ineligible and duplicate cases, analyzed 276 videos. Using a mDISCERN-based coding scheme,we assessed systematic content cues (e.g., whether videos provided evidence-based, balanced information and concrete self-management guidance) and heuristic account and production cues (e.g., provider type, on-screen text, graphics, and background music). We then modeled two outcomes: a quality index and a dissemination index based on likes, comments, and shares. Multiple regression with robust standard errors showed that systematic cues were strongly and positively associated with higher information quality, whereas heuristic cues had a smaller, marginally significant association. By contrast, neither systematic nor heuristic cues significantly predicted dissemination. Cross-equation tests indicated that the effects of both cue types differed significantly between the quality and dissemination models, pointing to a credibility–visibility gap: features that make diabetes videos more accurate and transparent do not automatically increase their early reach on Douyin.

Keywords: *short video, health communication, mDISCERN, dissemination, HSM.*

INTRODUCTION

With the rise of mobile internet and algorithmic recommendation, the dominant channel for health information has shifted rapidly from long-form text and video to short-video platforms. Short videos, by powerfully capturing attention and lowering cognitive load, enable fast diffusion and broad reach of health knowledge. Yet the same dynamics that amplify reach also produce high variance in information quality: rigorous, high-quality content is often submerged by entertaining formats, while emotional, sensational, oversimplified, or even misleading content may receive preferential algorithmic boosting and user engagement (Chen et al., 2025). This practical tension can be described as a visibility–credibility paradox: what attracts the most attention is not necessarily what should be trusted.In public health communication, the paradox bears directly on evidence-informed decision-making and behaviour change, and thus has substantial scholarly and practical importance.

*Corresponding author: emmamohamad@ukm.edu.my
E-ISSN: 2289-1528
https://doi.org/10.17576/JKMJC-2025-4104-13
Received: 6 November 2025 | Accepted: 5 February 2025 | Published: 12 December 2025

In China, Douyin has become a major channel through which patients and the general public encounter everyday health information, including advice on chronic conditions such as diabetes(Zhang et al., 2024). Many users turn to the platform for practical guidance on diet, medication, and exercise, yet existing evidence suggests that the quality of diabetes-related content is highly uneven, with some videos oversimplifying or even contradicting evidence-based guidelines. Because recommendation algorithms largely rely on engagement signals, emotionally appealing or sensational content may be promoted over rigorous but less eye-catching information (Song et al., 2024). For people who must manage diabetes over the long term, this visibility-first environment can shape risk perceptions and everyday decisions, underscoring the need to systematically examine how information quality and dissemination relate on Douyin.

Drawing on the heuristic–systematic model (HSM), we examine how two families of cues relate to both information quality and early dissemination on the platform. On the one hand are content-focused systematic cues, such as whether the video cites evidence, presents balanced information, and offers concrete guidance. On the other hand are heuristic cues tied to presentation and account characteristics, including provider type, on-screen text, graphics, and background music. We assess information quality using a coding index adapted from mDISCERN, and dissemination using a log-transformed Douyin Communication Index that combines likes, favourites, and shares with pre-specified weights. Rather than presuming that quality and reach move together, we model a "quality pathway" and a separate "dissemination pathway," applying the same set of cues in each. This setup allows us to ask whether the features that make diabetes videos more accurate and transparent also help them travel further on Douyin, or whether there are systematic tensions between credibility and visibility.

We analyzed 300 diabetes-related health videos retrieved from Douyin's "For You" feed and, after excluding ineligible and duplicate cases, obtained a final sample of 276 videos. Using parallel ordinary least squares models, we treated an mDISCERN-based index as the quality outcome and a log-transformed engagement index (logDCI1) as the dissemination outcome, and then used cross-equation tests to compare coefficients across pathways. Guided by HSM, the study addresses three research questions:

RQ1 (Quality pathway): Net of video duration (linear and quadratic terms), how are systematic cues and heuristic cues associated with content quality as assessed by mDISCERN?

RQ2 (Dissemination pathway): Net of video duration (linear and quadratic terms), how are systematic cues and heuristic cues associated with dissemination intensity, as measured by logDCI1?

RQ3 (Alignment vs. trade-offs): Do the effects of systematic and heuristic cues differ across the quality and dissemination pathways?

In this study, we define alignment as a pattern in which the coefficients for a given cue are statistically indistinguishable across the two models, and trade-offs as patterns in which the same cue shows coefficients that differ significantly in direction or magnitude across models. We test these differences using suest cross-equation equality tests.

RESEARCH BACKGROUND

The Heuristic–Systematic Model (HSM) posits two relatively independent yet concurrent routes of information processing: Systematic processing emphasises evidence, logic, source transparency, and argumentative strength, While heuristic processing relies on salient cues (e.g., on-camera presenter, on-screen text/animations, graphics, background music, account authority) to form "quick judgments" (Chu et al., 2021). Under dual-route (parallel) processing—where audiences may rely on low-effort heuristics while higher-effort systematic processing shapes rigour—the same cue may enhance quality without increasing dissemination (reach), and the reverse may also hold. Accordingly, systematic cues are more likely to elevate perceived expertise and credibility, whereas heuristic cues and presentational tempo are more likely to drive attention, dwell time, and interaction(Chu et al., 2021, Atkinson et al., 2023). This mapping provides a mechanistic framework for the visibility–credibility paradox.

Prior work typically examines one route at a time—using instruments such as mDISCERN/GQS to assess quality, or engagement metrics—likes, favourites, shares—or weighted composites such as the Douyin Communication Index (DCI) to measure dissemination (Jang et al., 2022a; Taştemur & Çiçek, 2023; Li et al., 2024; Arikan & Erol, 2025; Doğan & İpek, 2025; Şahin & Ay, 2025). Finally, engagement on short-video platforms is very uneven—many videos receive no likes or shares, while a small number attract very high counts. At the same time, prior studies have used a wide range of measures for dissemination, from raw counts to ratios and different composite indices. This inconsistency makes it difficult to compare studies directly, underscoring the need for more transparent and standardised operationalisations of dissemination (Ahmer et al., 2025).

To answer the above research questions, drawing on the HSM's parallel-processing assumptions, under a parallel-processing account, systematic cues can enhance information quality by increasing source transparency and evidentiary coherence. In contrast, heuristic cues act as processing scaffolds that support comprehension and help audiences keep track of the message. Accordingly, the study posits:

H1: Controlling for video duration, higher levels of systematic_sum are associated with higher mDISCERN scores.

H2: Under the same controls, higher levels of heuristic_sum are associated with higher mDISCERN scores.

Considering that platform ranking mechanisms favour cues that can be rapidly processed, heuristic elements are more likely to trigger early engagement and algorithmic amplification, whereas more systematic presentations invite deliberation and uncertainty resolution that may not immediately translate into interactions (Shin et al., 2022; Bird, 2024). On this basis, the study advance the following diffusion-path hypotheses:

H3: Controlling for video duration, systematic-sum will show a non-positive association with logDCI1.

H4: Controlling for video duration, higher heuristic-sum will be positively associated with logDCI1.

H5: The effect of the same cue on quality (mDISCERN) differs from its effect on diffusion (logDCI1), in magnitude and/or direction.

METHODOLOGY

a. Study Design and Analytic Framework

A content analysis combined with secondary analysis of platform engagement data was performed, using short video as the unit of analysis. Guided by the Heuristic–Systematic Model (HSM), this study decouple "content quality" and "dissemination intensity" into two parallel outcome pathways, estimate them separately, and then apply seemingly unrelated estimation (suest) to test cross-equation equality of identical predictors—thereby testing whether the same content cues affect information quality and dissemination in the same way, or whether a tension/trade-off between the two emerges.

Pathway 1 (Quality): dependent variable = mDISCERN total score.
Pathway 2 (Dissemination): dependent variable = logDCI1.

Each model includes a composite index of systematic cues (systematic_sum), a composite index of heuristic cues (heuristic_sum), and video duration entered as linear and quadratic terms. Both models are estimated by OLS with HC3-robust standard errors. To assess whether identical predictors operate similarly across models, a suest cross-equation tests of coefficient equality were performed, thereby evaluating alignment versus trade-offs between the quality and dissemination pathways.

b. Data Source and Sample Construction

To minimise personalised recommendation bias, all collection and screening were performed under a logged-out and private/incognito workflow: before each round, researchers cleared cache/cookies/history, opened an incognito window, did not log in, made no interactions (no follows, likes, comments, or long pauses), and loaded no platform-related extensions or scripts (Solomos et al., 2021). The researchers then ran in-platform searches on Douyin with the Chinese keyword "糖尿病" (diabetes) and related tags, checking results in the platform's default order.

The study included short videos that primarily covered diabetes education—knowledge, risk communication, lifestyle management, or treatment—and excluded videos aimed mainly at advertising/commerce, clearly irrelevant items, and re-edited mashups in which medical information was not the primary focus. All videos were verified one-by-one with links recorded; two researchers screened independently, resolving discrepancies by discussion.

We employed a cross-sectional design to systematically collect diabetes-related short videos on Douyin. Following prior research on health content on short-video platforms (Wang, 2025a; Wu et al., 2025), all searches were conducted while logged out to minimise bias from personalised recommendation. On three evenings—8, 15, and 22 September 2025 (UTC+8)—we opened the app without an account, entered the Chinese keyword "diabetes" (糖尿病), and retrieved videos from the default "For You" recommendation feed.

In total, 300 videos were initially retrieved across the three collection waves. After removing ineligible and duplicate cases, the final analytic sample comprised 276 unique diabetes-related videos. For the main analyses, we estimated two ordinary least squares (OLS) regression models with HC3 robust standard errors: a "quality pathway" model predicting an mDISCERN-based information quality index and a "dissemination pathway" model predicting a log-

E-ISSN: 2289-1528

https://doi.org/10.17576/JKMJC-2025-4104-13

transformed engagement index (logDCI1). Both models included two focal indices (systematic cues and heuristic cues) and several control variables (video length and its square).

From a statistical perspective, a sample of 276 cases is adequate for this design (Cao et al., 2025, Xu et al., 2025). Common rules of thumb for multiple regression recommend at least 10–15 observations per predictor; with fewer than ten predictors in each model, our case-to-predictor ratio comfortably exceeds these guidelines. In addition, previous content-analytic studies of health videos on short-video platforms have used comparable sample sizes (on the order of 100–300 videos), suggesting that our sample is well within the typical range for this area of research (Sun et al., 2023; Tu et al., 2025; Vitale et al., 2025; Wang, 2025b).

c. Measures and Operationalization

We assessed information quality using the mDISCERN instrument, which evaluates five aspects of health information: clarity of presentation, relevance to the topic, traceability of sources, robustness of the underlying evidence, and impartiality of the discussion. Each video received a total mDISCERN score ranging from 0 to 5, with higher scores indicating higher information quality (Jang et al., 2022b; Liu et al., 2024; Wu et al., 2025). Two trained coders independently rated all videos while blinded to each other's ratings in the initial round. Coding discrepancies were first discussed between the two coders, and when agreement could not be reached, a third researcher adjudicated the final score. Item wording and coding rules are provided in Appendix Table A1.

The researchers first computed a Douyin Communication Index (DCI) as a weighted sum of engagement: DCI=0.46×likes+0.37×favourites+0.17×shares (Chen & Liu, 2024). The weights (0.46/0.37/0.17) reflect the relative importance of these interactions for visibility on the platform (Pan & Chi, 2020; Liu et al., 2023; Yuan et al., 2024). To address skewness and zeros, this study applied a log transform: $logDCI1 = ln\, ln\, (DCI + 1)$. Because DCI was highly right-skewed with many zeros, this study applied a natural-log transform with a +1 offset (log1p), which reduces skewness, stabilises variance, and limits the influence of extreme values. Coefficients can thus be interpreted as associations with relative changes in DCI.

d. Key Predictors

We operationalised systematic cues as content features that signal more effortful, argument-based processing. Coders recorded the presence (1) or absence (0) of nine topical elements: prevention, disease knowledge, treatment and management, nutrition, exercise and rehabilitation, mental health, recent research, policy or public health measures, and case examples. These binary items were summed to create an equal-weight index (systematic_sum), with higher scores indicating that a video covered a broader set of systematic, content-focused cues. In line with the heuristic–systematic model, we treat these topical elements as parallel indicators of systematic processing, and there is no strong theoretical basis for assigning different numerical weights to individual topics. Using a simple sum therefore provides a transparent and conservative operationalisation of overall systematic content.

Heuristic cues captured account and presentation features that viewers may use as mental shortcuts when evaluating videos. We coded uploader type (government or professional organisation, news media, medical expert, commercial entity, or individual creator), shooting environment (medical vs. non-medical setting), follower-count category, the presence of an on-screen person (T_person, 0 = absent, 1 = present), on-screen text overlays (T_text, 0/1), graphics or visualisations (T_graphic, 0/1), overall presentation richness (combinations of person, text, and graphics), and background music (T_bgm, 0 = none, 1 = partial segments, 2 = throughout). Each feature was coded on a binary or ordinal scale and recoded so that larger values consistently indicated more of the cue. These items were then entered additively into a single index. The resulting heuristic_sum variable is the simple sum of all heuristic items, with higher scores indicating that a video contains more salient heuristic cues in terms of source, format, and production. Consistent with the HSM, we conceptualise these features as substitutable heuristic signals rather than components with a fixed theoretical ordering, which justifies treating them as equal-weight contributors to the overall heuristic cue index.

e. Controls

Video duration (in seconds) was coded from the platform display. To control for exposure and capacity effects—longer videos can both accrue more engagement and host more cues—we included both a linear term (length) and a quadratic term ($length^2$) in all models. This specification allows for potential nonlinearity (e.g., an inverted-U or diminishing-return pattern) in the relationship between length and the outcomes and permits calculation of the implied turning point. We treat duration as a pre-specified control to reduce bias rather than as a substantive focus of interpretation and therefore do not emphasise its coefficients (Li et al., 2025).

f. Coding Procedure and Reliability

Two trained coders independently performed the content analysis and mDISCERN ratings following a unified coding manual. After joint training and pilot coding, each video was coded independently by both coders, and any discrepancies were resolved through discussion. Intercoder agreement was assessed using Cohen's κ for 14 key items. Reliability was high overall (mean $\kappa = 0.902$, range 0.799–1.000). Using 0.70 as an acceptable and 0.80 as a good threshold, all items met $\kappa \geq 0.70$ and 92.9% met $\kappa \geq 0.80$. Full item-level coefficients are reported in Appendix Table A2.

For data preparation, we first imported all variables as text to avoid problems with local number formats. We then removed extra spaces and stray punctuation, harmonised category labels where needed, and converted the variables used in the analyses to numeric form. Basic checks ensured plausible ranges (e.g., non-negative counts; video_length_sec ≥ 0). We also created a quadratic term for duration ($length\_sec^2$) so that the regression models could allow for a non-linear relationship between length and the outcomes.

For dissemination, we constructed a Douyin Communication Index (DCI) following previous work: $DCI = 0.46 \times likes + 0.37 \times favourites + 0.17 \times shares$. These weights reflect Douyin's algorithmic emphasis on likes as the primary engagement signal, with favourites and shares contributing smaller but still meaningful weight to a video's visibility (陈冬玲, 2020; 张菲菲, 2021). Because the DCI distribution was highly right-skewed and contained many zeros, we used a log-transformed version, $logDCI1 = \ln(DCI + 1)$, in all analyses to reduce skewness and limit the influence of extreme values.

For the variables used in the two main regression models (information quality, logDCI1, cue indices, and video length), coding was essentially complete and no cases were dropped because of missing data. A small amount of item non-response occurred only on auxiliary variables. We attempted to apply Little's MCAR test to assess the missing-data mechanism, but the very low level of missingness led to insufficient degrees of freedom, so the test was not informative. Given the minimal and pattern-free missingness, we proceeded with complete-case analysis, which in practice retains all 276 videos in the main models.

g. Model Specification

We estimated two regression models corresponding to the quality (RQ1) and dissemination (RQ2)pathways. For information quality (RQ1), we used mDISCERN as a summed score that approximates a continuous outcome. We specified a linear model with video duration and its square:$mDISCERN_i = \alpha_1 + \beta_{1s}\, systematic\_sum_i + \beta_{1h}\, heuristic\_sum_i + \gamma_{11}\, length_i + \gamma_{12}\, length_i^2 + \varepsilon_{1i}$(Nwangburuka et al., 2023).

Ordinary least squares (OLS) with HC3-robust standard errors provides straightforward marginal-effect estimates and facilitates comparison with the dissemination model (Burton, 2021; Ibrahim, 2021; Newton et al., 2024). Including systematic_sum and heuristic_sum in the same equation reflects the dual-route logic of the heuristic–systematic model and allows us to test their separate associations with quality. Duration (length and length²) is treated as a pre-specified control to reduce confounding from exposure and capacity (longer videos can accrue more engagement and host more cues) and to allow for modest non-linearity (e.g., diminishing returns). Coefficients are interpreted as conditional associations: holding other variables constant, a positive and statistically significant coefficient on systematic_sum or heuristic_sum indicates that stronger cues are associated with higher mDISCERN.

The RQ2 (dissemination route) uses log(DCI_i+1) as the outcome, where DCI = 0.46·likes + 0.37·favourites + 0.17·shares and the log transform addresses zeros and right-skew. For comparability with the quality model, the study include the same predictors—sy stematic_sum and heuristic_sum—and control for duration and its square. Because longer videos can both accrue more engagement and host more cues and to allow for curvature in the length–dissemination relationship (e.g., an inverted-U/diminishing-return pattern) (Nwangburuka et al., 2023). When the data support such curvature, the study reports the turning point—the approximate duration at which predicted dissemination is highest—in seconds/minutes. Duration is used to purge bias rather than to make a substantive claim. The study do not emphasise its coefficients.

$$\begin{aligned} ln\, ln\, (\, DCI_i + 1) \\ &= \alpha_2 + \beta_{2s}\, systematic\_sum_i + \beta_{2h}\, heuristic\_sum_i + \gamma_{21}\, length_i \\ &+ \gamma_{22}\, length_i^2 + \varepsilon_{2i} \end{aligned}$$

Using the same predictors and controls in both models makes the two pathways directly comparable. In this log-linear form, coefficients can be read as approximate percentage changes in the DCI index for a one-unit change in each predictor, holding other variables constant, but our main focus is on the sign and significance of the cue indices rather than precise percentage interpretations.

To test whether the same cues operate similarly on the quality and dissemination pathways, we first estimated the two models separately by OLS with HC3-robust standard errors. We then used Stata's suest command to place both equations in a joint covariance framework and ran Wald tests of cross-equation coefficient equality. In particular, we tested whether the coefficients on systematic_sum and heuristic_sum were statistically identical across the two equations, and also considered joint tests that included duration terms. All tests were two-sided with $\alpha = 0.05$. We report χ^2 statistics and p-values, and interpret them alongside the point estimates and 95% confidence intervals. When equality is not rejected ($p \geq 0.05$), we treat the effects as aligned across outcomes; when equality is rejected ($p < 0.05$), we interpret the differences as evidence of trade-offs between quality and dissemination.

We examined Q–Q plots of residuals for both models and found only mild deviations from normality, which are mitigated by the use of HC3-robust standard errors. Predicted values with 95% confidence intervals were plotted over the p1–p99 range of video length to illustrate any non-linear duration patterns.(see Figures S1–S2). Nonlinearity interpretation: Based on model estimates, predicted lines with 95% CIs were plotted over the p1–p99 range of video length (main text Figures 1–2).

RESULTS

We estimated separate OLS models with HC3-robust standard errors for the quality pathway (dependent variable: mDISCERN) and the dissemination pathway (dependent variable: logDCI1) using the final analytic sample of N = 276 videos. The quality model achieved an R^2 of about 0.22, indicating that the cue indices and account controls explain a modest but meaningful share of the variance in information quality. The dissemination model achieved an R^2 of about 0.19, suggesting that early engagement on Douyin is partly, but not fully, accounted for by the same set of predictors. This pattern is consistent with the idea that diffusion intensity is also shaped by platform-side factors such as topic salience, account baselines, and posting windows. Full coefficients for both models are reported in Table 1. For transparency, model diagnostics (Q–Q plots) and a visualisation of duration effects are provided in Figures S1–S3.

a. Quality Route (RQ1: mDISCERN)

In the quality model, both cue indices are positively related to mDISCERN (Table 1). Net of provider type, shooting environment, follower category, and video length, the systematic cue index shows a statistically significant association with information quality ($b = 0.127$, $SE = 0.029$, $p < 0.001$, 95% CI [0.070, 0.185]). Substantively, videos that contain more structured, source-linked information and clearer risk communication receive higher quality scores.

The heuristic cue index is also positively associated with mDISCERN ($b = 0.625$, $SE = 0.303$, $p = 0.040$, 95% CI [0.029, 1.221]), although this effect is modest and should be interpreted cautiously. Videos featuring on-screen presenters, text overlays, graphics, and background music tend to be rated as higher quality, over and above their systematic content. Among the controls, provider type is negatively associated with quality ($b = -0.293$, $SE = 0.061$, $p < 0.001$), while follower category, environment, and length do not show clear associations. The model explains about 22% of the variance in mDISCERN ($R^2 = 0.216$).The Q–Q plot assesses the normality of regression residuals.

For the RQ1 model, the residuals lie close to the 45° line with only mild departures at the tails; the upper tail deviates slightly more than the lower, indicating modest heavy-tailedness/positive skew. This pattern is common in short-video data and does not threaten the consistency of OLS coefficient estimates. For inference, the study reports HC3 heteroskedasticity-robust standard errors, which make results less sensitive to minor non-normality and heteroskedasticity. Overall, the residual diagnostics do not pose a substantive threat to our conclusions; the study therefore interpret results on the basis of the coefficient estimates and their robust confidence intervals. The quality model's Q–Q plot (Figure 1) exhibits slight right-tail deviation from the 45°line—consistent with a few high-scoring cases—but overall supports the linear model assumptions.

Table 1: OLS regressions: Quality (mDISCERN) vs Dissemination (logDCI1)

Predictor	Quality pathway: mDISCERN	Dissemination pathway: logDCI1
Systematic cues	0.127*** (0.029)	−0.110 (0.097)
Heuristic cues	0.625* (0.303)	−1.936 †(0.989)
Provider type	−0.293*** (0.061)	0.050 (0.212)
Shooting environment	−0.055 (0.126)	−0.446 (0.427)
Follower category	0.000 (0.060)	1.191*** (0.190)
Video length	0.000 (0.000)	0.001 (0.001)
Constant	2.302*** (0.367)	6.953*** (1.167)
R^2	0.216	0.192
Adjusted R^2	0.199	0.174
F statistic	14.55	10.68
Observations	276	276

Note: Ordinary least squares (OLS) regressions with HC3 robust standard errors in parentheses. The quality outcome is mDISCERN; the dissemination outcome is log(DCI + 1) based on the Douyin Communication Index (DCI). Systematic and heuristic cue indices and control variables are defined in the Methods section.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$.

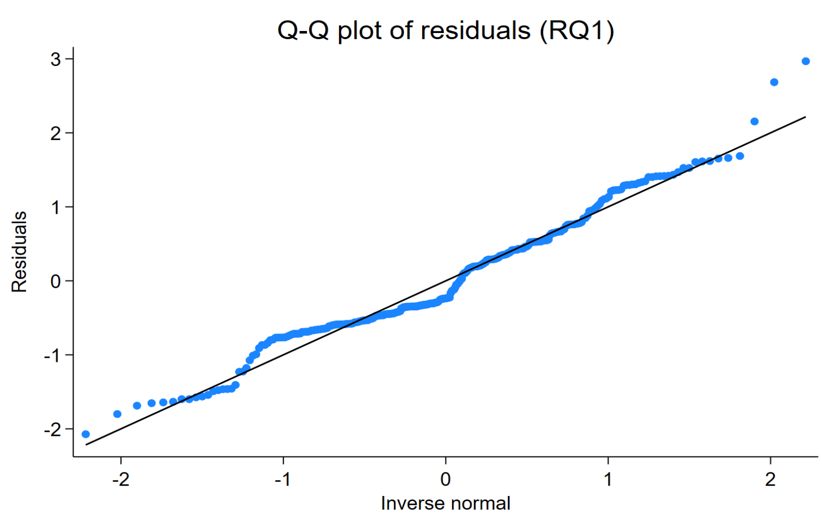

Figure 1: The quality model's Q–Q plot

https://doi.org/10.17576/JKMJC-2025-4104-13

b. *Diffusion Route (RQ2: logDCI1)*

The dissemination model paints a different picture. Net of the same controls, neither systematic nor heuristic cues significantly predict logDCI1, and both coefficients are negative (systematic cues: b = −0.110, SE = 0.097, p = 0.260; heuristic cues: b = −1.936, SE = 0.989, p = 0.051). Although the estimates are imprecise, their negative signs suggest a possible suppression pattern in which videos with more systematic or richly cued presentation do not obtain higher early engagement and may even attract slightly less, a pattern that warrants further investigation.

By contrast, follower category is strongly and positively associated with dissemination (b = 1.191, SE = 0.185, $p < 0.001$), indicating that videos from accounts with larger follower bases receive substantially more early engagement. Provider type, environment, and video length have small and statistically non-significant effects. Overall explanatory power for dissemination is modest ($R^2 = 0.192$), consistent with the idea that diffusion is driven by a mix of platform-side and network factors that extend beyond the coded content cues.

The Q–Q plot for the dissemination model (Figure 2) is close to the 45° line, with only a small deviation in the upper tail caused by a few very popular videos. With HC3 robust standard errors, these minor departures from normality are unlikely to affect our conclusions.

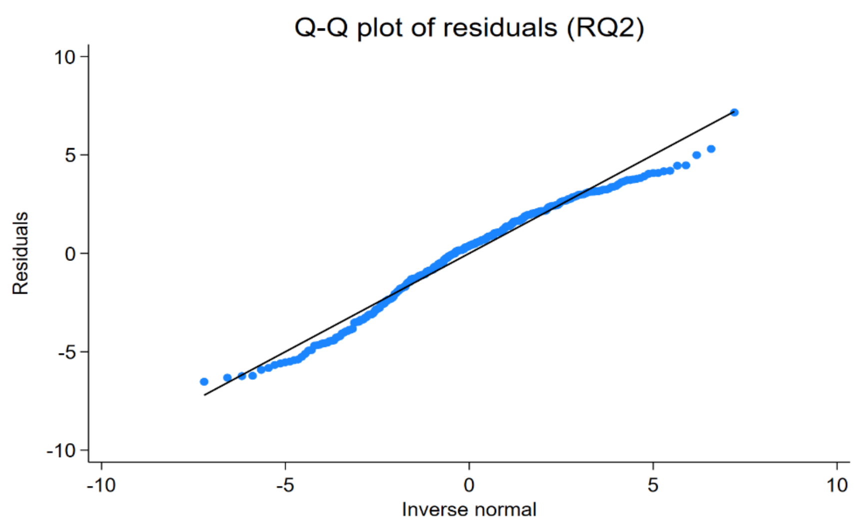

Figure 2: logDCI1 model's Q–Q plot

Figure S3 ("Predicted mDISCERN over length") based on the RQ1 regression, shows the adjusted relationship between video length and mDISCERN holding other variables constant. The central prediction line is nearly flat (slightly downward), indicating that within the observed range the marginal effect of length on quality is minimal. The 95% confidence band is wide and overlaps a zero effect throughout, suggesting the trend is not statistically significant. The band widens markedly at longer durations, reflecting sparser observations and greater predictive uncertainty in that region. Overall, the pattern aligns with the regression table: video length has no clear effect on quality, so we focus on the conditional effects of systematic and heuristic cues.

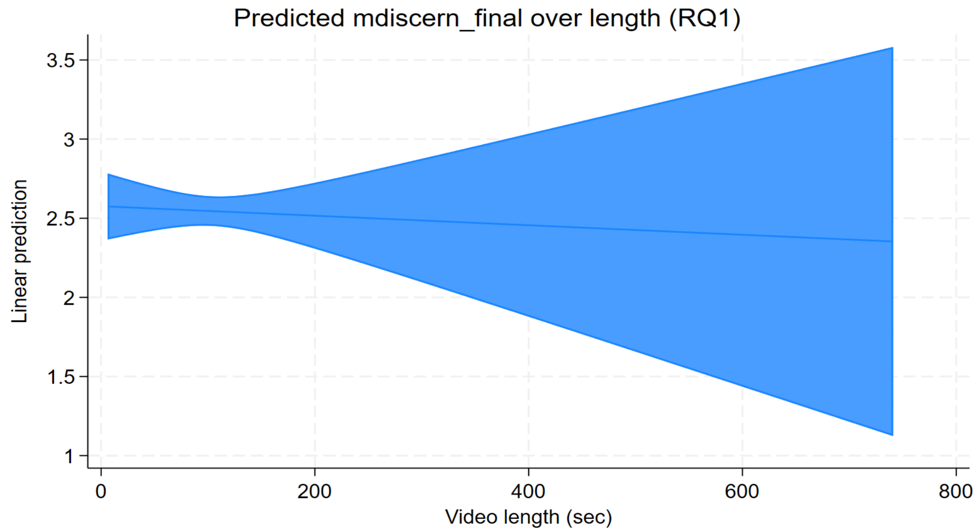

Figure 3: Predicted mDISCERN over length plots

DISCUSSION

a. *Who Improves mDISCERN?*

Within the quality pathway (DV = mDISCERN), and controlling for provider type, shooting environment, follower category, and video length, the composite index of systematic cues shows a clear positive association with information quality (Table 1; b = 0.127, SE = 0.029, $p < 0.001$, R^2 = 0.216). Videos that provide more structured explanations, cite sources, disclose risks, and align conclusions with evidence receive higher scores on the mDISCERN scale.

The composite index of heuristic cues is also positive (b = 0.625, SE = 0.303, p = 0.040), although this effect is only marginally significant and should be interpreted with some caution. On-screen presenters, captions or text overlays, graphics, and background music appear to go together with higher perceived quality rather than undermining it. In a fast, attention-limited short-video environment, these elements may plausibly help viewers follow the message—for example by highlighting key points and maintaining attention—even though we did not directly measure cognitive processing or cognitive load.

Among the control variables, some provider types score significantly lower than the reference group, suggesting that who posts the video matters for rated quality. Follower category, environment, and video length show no strong independent association with mDISCERN once the cue indices are included. Overall, the quality model supports the idea that content-related systematic cues are the primary drivers of perceived informational quality, with heuristic cues providing a smaller, more uncertain contribution.

b. *What Determines logDCI1?*

The dissemination pathway (DV = logDCI1) tells a different story. Net of the same controls, coefficients on both systematic and heuristic cues are negative and not conventionally significant (systematic_sum: b = −0.110, SE = 0.097, p = 0.260; heuristic_sum: b = −1.936, SE = 0.989, p = 0.051, R^2 = 0.192). Together, these estimates suggest that adding more systematic or heuristic cues does not translate into higher early engagement on Douyin; if anything, the point estimates hint at a small negative association that remains imprecisely estimated.

By contrast, follower category has a large and highly significant positive coefficient ($b = 1.191$, $SE = 0.185$, $p < 0.001$), indicating that videos from accounts with larger follower bases receive substantially more likes, favourites, and shares in the early phase. Provider type, environment, and length show only weak or inconsistent relationships with logDCI1 once follower base is taken into account.

These patterns are consistent with the idea that diffusion intensity—operationalised as a weighted, log-transformed index of likes, favourites, and shares—is shaped less by the specific content cues coded here and more by platform and network factors such as account endowments, recommendation dynamics, timing, and unmeasured affective features (e.g., humour, controversy, surprise).

c. Are Quality and Diffusion Driven by the Same Factors?

To test whether the same cues operate similarly across the two pathways, we used seemingly unrelated estimation (suest) and cross-equation Wald tests. For systematic cues, the equality constraint across equations is rejected ($\chi^2(1) = 6.24$, $p = 0.012$), indicating that the positive effect on mDISCERN is significantly stronger than the (small and negative) effect on logDCI1. For heuristic cues, equality is also rejected ($\chi^2(1) = 6.73$, $p = 0.010$): heuristic_sum is positively and marginally associated with quality, but negatively and marginally associated with dissemination.

At minimum, these tests show that perceived quality and early reach are not driven by the cue indices in the same way. For systematic cues, the findings support a “visibility–credibility gap”: features that raise verifiability and structure contribute to higher quality evaluations but do not confer clear advantages in early algorithmic exposure. For heuristic cues, the positive association with quality and the negative association with dissemination point to a possible trade-off pattern, although the dissemination effect is estimated with greater uncertainty.

These results imply that improving quality alone is unlikely to automatically boost early reach. Instead, creators and public-health institutions may need to pair strong systematic content and thoughtful presentation with explicit distribution strategies—such as timing, serialisation, metadata choices, and cross-platform promotion—rather than expecting high-quality production to “spread itself.”

CONTRIBUTIONS

a. *Theoretical and M*ethodological *Contributions*

This study adapts the heuristic–systematic model to short-form video by explicitly separating and jointly modelling a quality pathway (mDISCERN) and a dissemination pathway (logDCI1) for the same set of diabetes videos on Douyin. Using common cue indices in both models and formally testing cross-equation equality, we show that perceived information quality and early algorithmic reach respond differently to the same content features, clarifying a visibility–credibility gap in health short videos. The findings also nuance the role of heuristic cues: rather than treating them purely as threats to rigour, our results suggest that presentation elements can coexist with, and modestly support, higher quality evaluations, even if they do not reliably boost early diffusion.

Methodologically, the study offers an operationalised coding scheme for systematic and heuristic cues that can be reused in future research, together with outcome measures that are both verifiable and aligned with platform practice (mDISCERN for quality and a log-transformed DCI for dissemination). The combination of dual-pathway modelling and cross-equation tests via suest provides a template for examining alignment and trade-offs between quality and reach in other digital health settings.

b. Empirical and Practical Contributions

Empirically, the study provides within-platform, dual-route evidence from 276 diabetes-related videos on Douyin. Systematic cues consistently predict higher mDISCERN scores, and heuristic cues show a smaller, more uncertain positive association with quality, whereas neither cue index reliably increases early dissemination once follower base is taken into account. Follower category emerges as a strong predictor of engagement, underscoring the influence of account endowments and platform dynamics alongside content cues. These results caution against assuming that credibility-enhancing features will automatically be rewarded by the recommendation system.

Practically, the findings translate into a dual-track approach for public-health short videos. A “quality stack” grounded in systematic cues—clear structure, explicit sourcing and evidence, risk and uncertainty disclosures, and conclusion–evidence consistency—can be complemented by a “reach stack” that combines appropriate heuristic presentation with deliberate distribution strategies (timing, serial formats, hooks, early interaction prompts, and cross-platform seeding). For platforms, the documented gap between quality and early reach highlights the potential value of quality-aware ranking and visible quality labels in mitigating the amplification of high-reach, low-quality health content.

LIMITATIONS

Several limitations should be noted. First, the data come from 276 diabetes-related videos collected from Douyin over three logged-out waves within a single month. Generalisation to other health conditions, longer time windows, and other platforms with different audiences and recommendation rules remains to be established. Second, the design is cross-sectional and engagement was frozen at the time of capture to avoid post-scrape growth bias. As a result, we cannot speak to causal dynamics or to later diffusion bursts that may occur after the initial exposure window.

Third, the dissemination model relies on content features only and omits account- and network-level variables (e.g., follower base, prior engagement history) and direct measures of algorithmic placement, which likely explain a substantial share of residual variance in logDCI1. Fourth, our heuristic composite focuses on information-design cues (presenter, captions, graphics, background music) and does not directly encode affective or interpersonal heuristics such as emotional intensity, humour or controversy, and parasocial cues, which may be important drivers of diffusion. Finally, although coding followed a structured protocol and intercoder agreement was high, content analysis remains subject to human error and judgement.

Future work could extend the present design in several directions: incorporating affective and interpersonal heuristics into the cue indices; adding account- and network-level covariates to better explain dissemination; and strengthening causal inference through longitudinal designs or quasi-experiments (e.g., around policy or interface changes). Cross-platform and cross-topic replications would help distinguish general from domain-specific patterns in the relationship between quality cues and reach.

CONCLUSION

This study examined how systematic and heuristic cues in short-form health videos relate to information quality and early dissemination on Douyin. Systematic cues robustly improve mDISCERN scores, and heuristic cues show a smaller, more uncertain positive association with quality, but neither cue index reliably boosts early engagement once follower base is controlled. Cross-equation tests confirm that the same cues operate differently across the quality and dissemination pathways, quantifying a gap between credibility and visibility in this setting.

These findings suggest that systematic cues should remain the backbone of credible health communication on short-video platforms, while heuristic cues can be used to support comprehension and attention rather than as substitutes for substance. At the same time, reach is strongly shaped by account endowments and distribution mechanics, not only by production value. For creators and public-health agencies, effective practice likely requires pairing a quality stack (grounded in systematic cues) with a reach stack that combines thoughtful presentation and distribution strategies. For platforms, incorporating quality signals into ranking—through verified-source badges, structured quality fields, or visible quality labels—could help reduce the systemic imbalance between high-reach and high-quality content in sensitive health domains.

The present evidence comes from one topic, one platform, and one time window, and is based on observational data. As short-video platforms continue to evolve, future research combining content analysis with account- and network-level data, longitudinal designs, and experimental or quasi-experimental approaches will be important for clarifying when and how quality can be translated into reach in digital health communication.

BIODATA

Shan Chen is a PhD candidate at the Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia. Her work focuses on health communication and social media. Email: p126054@siswa.ukm.edu.my

Xi Xi Zhao is a PhD candidate at the Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia. Email: p132328@siswa.ukm.edu.my

Assoc. Prof. Dr Emma Mirza Wati Mohamad is an associate professor in the Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia. Her research covers health communication and social behaviour change communication. Email: emmamohamad@ukm.edu.my

Dr. Arina Anis binti Azlan is a lecturer in the Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia. Her research covers health communication, information management, and communication behaviour. Email: arina@ukm.edu.my

REFERENCES

Ahmer, A., Williams, A., Antonacci, G., Harris, M., & Irvine, R. (2025). *Improving acceptability of mHealth apps—The use of the technology acceptance model to assess the acceptability of mHealth apps: Systematic review.* PROSPERO 2024. https://www.crd.york.ac.uk/PROSPERO/view/CRD42024532974

Arikan, H., & Erol, E. (2025). *Quality and reliability evaluation of YouTube® exercises content for temporomandibular disorders*. *BMC Oral Health, 25*(1), 301.

Atkinson, T., Brown, E., Jones, G., Sage, K., & Wang, X. (2023). "I assumed it would be somebody who had a stroke that was doing this": Views of stroke survivors, caregivers, and health professionals on tailoring a relaxation and mindfulness intervention. *Healthcare, 11*(3): 399.

Bird, Y. (2024). Rating anticipation and strategic downward selection in consumer-generated rating systems: Evidence from a peer-to-peer platform market. *Journal of Management Studies.*

Burton, A. L. (2021). OLS (Linear) regression. In J. C. Barnes, & D. R. Forde (Eds.), *The Encyclopedia of Research Methods in Criminology and Criminal Justice* (Chap. 104, pp. 509-514). Wiley.

Cao, J., Zhang, F., Zhu, Z., & Xiong, W. (2025). Quality of cataract-related videos on TikTok and its influencing factors: A cross-sectional study. *Digital Health, 11*, 20552076251365086.

Chen, M., Lin, X., Zhou, R., & Fan, G. (2025). U-shaped association between social media usage frequency and suggestibility by internet health information in Chinese online population with pre-diabetes and diabetes: A cross-sectional study. *BMC Public Health, 25*(1), 1525.

Chen, X., & Liu, Y. (2024). Chinese libraries' communication influence based on the Douyin communication index. *Library Hi Tech*.

Chu, X., Liu, Y., & Chen, X. (2021). The impacts of content and source factors on consumers' liking toward advertisements: An HSM-based framework. *2020 International Conference on Applications and Techniques in Cyber Intelligence: Applications and Techniques in Cyber Intelligence* (ATCI 2020) (pp. 106-112). Springer.

Doğan, G., & İpek, H. (2025). Evaluating the quality and reliability of YouTube videos on Hirschsprung's disease: A comprehensive analysis for patients, parents and health professionals. *Pediatric Surgery International, 41*(1), 173.

Ibrahim, N. (2021). Dutch disease effects in the Azerbaijan economy: Results of multivariate linear ordinary least squares (OLS) estimations. *Экономический журнал Высшей школы экономики, 25*(2), 309-346.

Jang, C. W., Bang, M., Park, J. H., & Cho, H. E. (2022a). Value of online videos as a shoulder injection training tool for physicians and usability of current video evaluation tools. *International Journal of Environmental Research and Public Health, 19*(22), 15177.

Jang, C. W., Kim, M., Kang, S. W., & Cho, H. E. (2022). Reliability, quality, and educational suitability of TikTok videos as a source of information about scoliosis exercises: A cross-sectional study. *Healthcare, 10*(9), 1622.

Li, Q., Cicirelli, F., Vinci, A., Guerrieri, A., Qi, W., & Fortino, G. (2025). Quadruped robots: Bridging mechanical design, control, and applications. *Robotics, 14*(5), 57.

Li, Z., Yan, C., Lyu, X., Li, F., & Zeng, R. (2024). Assessing quality and reliability of online videos on tachycardia: A YouTube video-based study. *BMC Public Health, 24*(1), 2620.

Liu, H., Peng, J., Li, L., Deng, A., Huang, X., Yin, G., ... & Liang, Y. (2024). Assessment of the reliability and quality of breast cancer related videos on TikTok and Bilibili: Cross-sectional study in China. *Frontiers in Public Health, 11*, 1296386.

Liu, Y., Chiu, D. K., & Ho, K. K. (2023). Short-form videos for public library marketing: Performance analytics of Douyin in China. *Applied Sciences, 13*(6), 3386.

Newton, I. H., Hasan, M. H., Razzaque, S., & Roy, S. K. (2024). Assessment of climate-induced rice yield using ordinary least squares (OLS) regression analysis: A case study from coastal context. *Earth Systems and Environment, 8*(4), 1437-1451.

Nwangburuka, C., Ijomah, M. A., & Nwakuya, M. T. (2023). Heteroscedasticity of unknown form: A comparison of five heteroscedasticity-consistent covariance matrix (hccm) estimators. *Global Journal of Pure and Applied Sciences, 29*(1), 83-90.

Pan, C., & Chi, R. (2020). Analysis and research on operation of Tik Tok accounts of Chinese airlines. *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology* (ICCASIT) (pp. 936-939). IEEE.

Şahin, H., & Ay, B. K. (2025). YouTube as a source of information on pulmonary rehabilitation for COPD patients. *Duzce Medical Journal*.

Shin, D., Zaid, B., Biocca, F., & Rasul, A. (2022). In platforms we trust? Unlocking the black-box of news algorithms through interpretable AI. *Journal of Broadcasting & Electronic Media, 66*(2), 235-256.

Solomos, K., Kristoff, J., Kanich, C., & Polakis, J. (2021). *Tales of favicons and caches: Persistent tracking in modern browsers*. Network and Distributed System Security Symposium.

Song, G., Wang, Y., Chen, X., Hu, H., & Liu, F. (2024). Evaluating user engagement in online news: A deep learning approach based on attractiveness and multiple features. *Systems, 12*(8), 274.

Sun, F., Zheng, S., & Wu, J. (2023). Quality of information in gallstone disease videos on TikTok: Cross-sectional study. *Journal of Medical Internet Research, 25*, e39162.

Taştemur, M., & Çiçek, C. M. (2023). Quality analysis of YouTube videos in the management of hyperlipidemia in adults. *Journal of Health Sciences and Medicine, 6*(5), 1074-1079.

Tu, J., Zhang, C., Zhang, H., Liang, L., & He, J. (2025). Evaluating the reliability and quality of knee osteoarthritis educational content on TikTok and Bilibili: A cross-sectional content analysis. *Digital Health, 11*, 20552076251366390.

Vitale, S. G., Angioni, S., Saponara, S., Sicilia, G., Etrusco, A., D'Alterio, M. N., ... & Riemma, G. (2025). TikTok as a platform for hysteroscopy information: An analytical video-based cross-sectional study to assess quality, reliability, and accuracy. *International Journal of Gynecology & Obstetrics, 168*(1), 353-361.

Wang, H., Zhang, H., Cao, J., Zhang, F., & Xiong, W. (2025a). Quality and content evaluation of thyroid eye disease treatment information on TikTok and Bilibili. *Scientific Reports, 15*(1), 25134.

Wang, K., Tan, X., & Liu, P. (2025b). Quality and reliability of Chinese short videos on TikTok related to chronic renal failure: Cross-sectional study. *Frontiers in Public Health, 13*, 1652579.

Wu, J., Wu, G., Che, X., & Guo, J. (2025). The quality and reliability of short videos about hypertension on TikTok: A cross-sectional study. *Scientific Reports, 15*(1), 25042.

Xu, R., Ren, Y., Li, X., Su, L., & Su, J. (2025). The quality and reliability of short videos about premature ovarian failure on Bilibili and TikTok: Cross-sectional study. *Digital Health, 11*, 20552076251351077.

Yuan, Y., Li, Y., & Sun, H. (2024). Utilizing multidimensional features to predict the dissemination-force of emergency short videos. *Proceedings of the 24th ACM/IEEE Joint Conference on Digital Libraries*, 1-11.

Zhang, B., Kalampakorn, S., Powwattana, A., Sillabutra, J., & Liu, G. (2024). Oral diabetes medication videos on Douyin: Analysis of information quality and user comment attitudes. *JMIR Formative Research, 8,* e57720-e57720.

陈冬玲. (2020). 基于DCI指数的我国图书馆抖音平台传播策略研究. *图书馆研究与工作, 12,* 35-39.

张菲菲. (2021). HPV健康短视频信息采纳与劝服效果实证研究.