

Turkish YouTube Videos as a Source of Information on Biologic Therapies in Rheumatic Diseases: An Assessment of Quality, Reliability, and Content

Nihal Lermi, Zeynep Fural

Department of Rheumatology, University of Health Sciences, Istanbul Kartal Dr. Lütfi Kırdar Health Research Center, Istanbul, Türkiye

ORCID  of the author(s)

NL: <https://orcid.org/0000-0002-2774-5913>
ZF: <https://orcid.org/0000-0001-5859-2502>

Corresponding Author

Nihal Lermi

University of Health Sciences, Istanbul Kartal Dr. Lütfi Kırdar Health Research Center, Department of Rheumatology, Istanbul, Türkiye
E-mail: nihalyilmaz16@hotmail.com

Ethical Approval

Ethics committee approval was not required for this study because it involved publicly available YouTube videos and did not include human participants or identifiable personal data. The Declaration of Helsinki was not applicable because no human participants were involved.

Informed Consent

Not applicable.

Conflict of Interest

No conflict of interest was declared by the authors.

Financial Disclosure

The authors declared that this study received no financial support.

CRediT Author Statement

Nihal Lermi: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Supervision.
Zeynep Fural: Investigation, Writing – review & editing.

Use of AI Statement

AI-assisted tools were used to support English language editing, improve the clarity and readability of the manuscript, and assist with table formatting and presentation. All AI-assisted modifications were reviewed and verified by the authors, who take full responsibility for the accuracy and integrity of the manuscript.

Data Availability

The dataset generated and analyzed during the current study is available from the corresponding author upon reasonable request.

Published

2026 October 17

Copyright 2026 © The Author(s)



This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

<https://creativecommons.org/licenses/by-nc-nd/4.0/>



Abstract

Background/Aim: YouTube is increasingly used as a source of health-related information; however, the quality and reliability of Turkish-language content on biologic therapies in rheumatic diseases remain unclear. This study aimed to evaluate the quality, reliability, usefulness, and content completeness of Turkish-language YouTube videos on biologic therapies.

Methods: In this cross-sectional study, 24 eligible videos were independently evaluated by two rheumatologists using the modified DISCERN (mDISCERN), Global Quality Score (GQS), usefulness classification, and an 11-item content checklist.

Results: Twenty videos (83.3%) were physician-presented. The median content score was 5 (range, 1–10), and no video addressed all 11 domains. Indications for biologic therapy were the most frequently covered domain (79.2%), whereas pregnancy/breastfeeding was the least frequently addressed (4.2%). The median mDISCERN and GQS scores were 2 (range, 1–4) and 3 (range, 1–5), respectively, and 21 videos (87.5%) were classified as useful. GQS correlated with mDISCERN ($\rho = 0.452$, $P = 0.026$) and content score ($\rho = 0.436$, $P = 0.033$). Video duration correlated strongly with content score ($\rho = 0.801$, $P < 0.001$) and moderately with GQS ($\rho = 0.474$, $P = 0.019$), whereas views and likes were not associated with quality, reliability, or content completeness.

Conclusion: Turkish-language YouTube videos on biologic therapies were generally useful but showed limited reliability and incomplete content. The marked lack of information on pregnancy and breastfeeding highlights an important gap in patient-oriented educational content.

Keywords: biologic therapy, patient education, rheumatic diseases, reliability, youtube

Introduction

The Internet has become one of the principal sources of health-related information for patients, and among the many platforms available online, YouTube has emerged as one of the most influential video-sharing sites for the dissemination of medical content, with hundreds of hours of new material uploaded every minute [1, 2]. Its free access, ease of use, and capacity to convey complex information through visual and narrative formats have made it an increasingly common resource for patients seeking to understand their diagnosis and treatment, often as a supplement to, rather than a substitute for, information obtained during clinical encounters [3, 4].

Biologic therapies have transformed the management of numerous rheumatic diseases, and patients with high disease activity despite conventional treatment are increasingly offered these agents [5]. Nevertheless, considerable uncertainty persists among patients regarding response variability, administration, and the long-term safety of biologic agents [3]. Because consultation time is limited, video-based platforms are increasingly used by patients to supplement counseling on issues such as the mechanism of action, route and frequency of administration, pre-treatment infection screening, adverse events, monitoring during treatment, and safety considerations in special populations, including pregnancy and breastfeeding [5, 6].

However, YouTube content is not subject to independent editorial or peer review before publication, and the diversity of its contributors, together with the absence of a formal quality-control mechanism, creates a considerable risk that patients will encounter incomplete, outdated, or misleading information [3, 7, 8]. A growing number of studies have accordingly evaluated the quality and reliability of YouTube videos on rheumatological topics—including secukinumab [5], the side effects of biologic therapy [3], musculoskeletal ultrasound [2], anakinra self-injection [9], self-administered subcutaneous anti-TNF injections [10], spondyloarthritis [11], systemic lupus erythematosus [12], and Sjögren's syndrome [13]—generally using validated instruments such as the Global Quality Score (GQS) [14] and a modified DISCERN (mDISCERN) tool adapted from the original written-health-information instrument of Charnock et al. [15]. These studies have consistently reported wide variability in video quality and reliability, and several have shown that videos uploaded or presented by physicians and professional/academic organizations tend to be of higher quality and reliability than those uploaded by patients or independent users [4, 5, 6, 16].

Most of this literature has focused on English-language videos, and Turkish-language YouTube content specific to rheumatology remains comparatively underexplored. Turkish-language videos have been examined for ankylosing spondylitis [6] and, outside rheumatology, for immunotherapy in oncology [17], both of which found that content produced by specialists was of substantially higher quality and reliability than content from non-expert or unidentified sources. However, to our knowledge, no prior study has specifically evaluated Turkish-language YouTube videos addressing biologic therapies used across the spectrum of rheumatic diseases. In addition, whereas most previous studies in this field have concentrated on overall quality and reliability scores, comparatively few have systematically

assessed which specific clinical content domains are actually covered within individual videos, and none, to our knowledge, has done so for biologic therapies in rheumatology using a dedicated, multi-domain checklist.

The present study was therefore designed to evaluate the quality, reliability, and content of Turkish-language YouTube videos providing information on biologic therapies in rheumatic diseases. In addition to the GQS and mDISCERN instruments used in prior research, we developed an 11-item content checklist specifically to assess the completeness of information on biologic therapies, spanning domains from indications and mechanism of action to administration, infection screening, monitoring, adverse events, and pregnancy/breastfeeding. We further examined the relationships between video popularity and engagement, presenter type, video duration, and these quality, reliability, and content measures.

Materials and methods

Study Design and Search Strategy

This cross-sectional study evaluated the quality, reliability, and content of Turkish-language YouTube videos providing information on biologic therapies in rheumatic diseases. As the study involved only publicly available online content and did not include human participants or personal data, ethics committee approval was not required. No identifiable user comments, usernames, or other personal information were collected or analyzed; only aggregate comment counts were recorded.

The YouTube searches were conducted between the evening of August 19 and the early morning of August 20, 2026. YouTube searches were performed using Microsoft Edge in InPrivate browsing mode without logging into a personal account to minimize the potential influence of previous search history and personalized recommendations. All searches were conducted using the same search strategy. No predefined limit was applied to the number of videos screened for each search term. For each search, the results page was continuously scrolled until no additional results were displayed, and all videos retrieved under the selected search settings were screened. The ‘Popularity’ sorting option was selected to prioritize videos with greater user visibility, consistent with previous YouTube-based health information studies that prioritized highly viewed content to capture videos with greater potential public impact [1, 5].

Initially, six combinations of Turkish search terms related to rheumatic diseases and biologic therapies were used: “romatolojik hastalıklar biyolojik tedavi,” “romatizmal hastalıklar biyolojik tedavi,” “romatizma biyolojik tedavi,” “romatolojik hastalıklar biyolojik ajanlar,” “romatizmal hastalıklar biyolojik ajanlar,” and “romatizma biyolojik ajanlar.” During pilot screening of the initial search results, the initial search terms yielded a limited number of relevant videos, and it was observed that relevant Turkish-language content frequently used the term “anti-TNF” rather than the broader terms “biologic therapy” or “biologic agents.” Therefore, to improve the comprehensiveness of the search, three supplementary anti-TNF searches were added. This refinement was introduced after inspection of the pilot search results but before formal video screening and evaluation. The initial six searches yielded 17, 13, 14, 4, 3, and 4 results,

respectively, corresponding to 55 search results before duplicate removal. Duplicate videos were identified using their unique YouTube video IDs, leaving 22 unique videos.

The supplementary searches used the terms “romatolojik hastalıklar anti-TNF,” “romatizmal hastalıklar anti-TNF,” and “romatizma anti-TNF.” These searches yielded 31, 26, and 34 results, respectively. After duplicate removal, 52 unique videos were identified from the anti-TNF searches. Most of these videos were not in Turkish. Nine additional Turkish-language candidate videos not represented in the initial search set were identified, of which eight met the eligibility criteria and were included.

After removal of duplicates and application of the eligibility criteria, 24 videos constituted the final study sample.

Eligibility Criteria

Videos were eligible if they were in Turkish and contained substantive information on biologic therapy in rheumatic diseases. Videos in which biologic therapy constituted the main topic were included. Videos addressing a broader rheumatologic topic were also considered eligible when biologic therapy was discussed in sufficient detail to allow meaningful assessment.

Videos were excluded if they were not in Turkish, were duplicates, referred to biologic therapy only incidentally without providing substantive information, focused on biologic drugs without sufficient relevance to rheumatic diseases, or consisted primarily of unrelated or promotional content. Patient-generated videos were not excluded solely on the basis of their source, as patient experiences and opinions regarding biologic therapies may also represent health information encountered by YouTube users.

Video Characteristics

For each included video, the video title, URL, upload date, duration, number of views, number of likes, number of comments, and presenter type were recorded. Videos were categorized according to the person presenting the information as physician-presented or patient/individual-user-presented.

Video characteristics were recorded independently of the subjective quality assessments and were therefore not considered rater-dependent variables. For videos in which comments were disabled, the number of comments was considered unavailable rather than being recorded as zero.

Content Assessment

To evaluate the completeness of information specifically related to biologic therapies in rheumatic diseases, a study-specific 11-item content assessment checklist was developed before video evaluation. The domains were selected by the two rheumatologist investigators based on clinically relevant information considered important for patients receiving biologic therapies and on content domains evaluated in previous YouTube-based studies of rheumatic diseases and biologic treatments [1, 3, 5, 9, 10]. The checklist was intended as a predefined framework for assessing the breadth of biologic-therapy information provided in each video rather than as a validated measurement instrument. The checklist evaluated the presence of the following information: (1) definition/general characteristics of biologic therapy; (2) mechanism of action; (3) indications/rheumatic diseases in which biologic therapies are used; (4) criteria for treatment initiation/patient selection; (5) route and/or frequency of administration; (6) pre-treatment infection screening/evaluation

and/or vaccination recommendations; (7) risk of infection; (8) other adverse events; (9) contraindications, precautions, and/or concomitant use with other rheumatologic treatments; (10) monitoring/follow-up during treatment; and (11) pregnancy/breastfeeding.

Each item was scored as present (1) or absent (0). The individual scores were summed to obtain a total content score ranging from 0 to 11, with higher scores indicating more comprehensive coverage of information related to biologic therapy.

Reliability and Quality Assessment

The reliability of the videos was assessed using the modified DISCERN (mDISCERN) instrument. The instrument consists of five yes/no questions. Each positive response receives 1 point and each negative response 0 points, resulting in a total score ranging from 0 to 5, with higher scores indicating greater reliability.

Overall video quality was evaluated using the Global Quality Score (GQS), a five-point scale ranging from 1 (poor quality) to 5 (excellent quality). GQS was assessed according to the overall flow, clarity, and usefulness of the information within the intended scope of each video rather than solely according to the number of content domains addressed. Consistent with previously used GQS categorizations in YouTube-based health information studies [16], scores of 1–2 were categorized as low quality, 3 as moderate quality, and 4–5 as high quality.

Usefulness Assessment

Videos were additionally classified according to their overall usefulness into one of four mutually exclusive categories, adapted from classifications used in previous YouTube-based health information studies [10, 16]: (1) useful medical information, defined as accurate and informative medical content; (2) misleading medical information, defined as inaccurate, incomplete, or potentially misleading medical content; (3) useful patient experience/opinion, defined as personal experience or opinion that provided relevant and potentially helpful information from the patient perspective; and (4) misleading patient experience/opinion, defined as personal experience or opinion containing inaccurate or potentially misleading medical claims. Videos containing both useful and misleading medical information were classified in the misleading category.

Evaluation Process and Inter-rater Agreement

All eligible videos were independently evaluated by two rheumatologists using the same standardized assessment form. The reviewers independently assessed the 11 content items, the five mDISCERN items, GQS, and usefulness without access to each other's ratings.

After completion of the independent assessments, all disagreements were reviewed and resolved by consensus. Consensus ratings were used for the final analyses, whereas the original independent ratings were retained for assessment of inter-rater reliability.

For the 11-item content checklist and the five individual mDISCERN items, inter-rater agreement was assessed using percentage agreement and pooled Cohen's kappa coefficients, calculated by combining item-level ratings across all videos. To account for the clustering of multiple item ratings within the same video, 95% confidence intervals (CIs) for the pooled kappa

estimates were obtained using cluster bootstrap resampling at the video level. Agreement for the total mDISCERN score was assessed using a two-way random-effects, single-measure, absolute-agreement intraclass correlation coefficient [ICC(2,1)]. Agreement for GQS was assessed using quadratic-weighted Cohen's kappa, and agreement for usefulness classification was assessed using unweighted Cohen's kappa. Ninety-five percent CIs were reported for the ICC and kappa estimates.

Statistical Analysis

Continuous variables were summarized using medians, interquartile ranges (IQRs), and minimum–maximum values, as appropriate, and categorical variables were presented as numbers and percentages. Frequencies of the individual content domains were calculated, and the total content score was determined for each video. Given the small sample size and distributional characteristics of the data, non-parametric tests were used for group comparisons. Physician-presented and patient/individual-user-presented videos were compared using the Mann–Whitney U test for continuous and ordinal variables. Given the small and markedly unequal group sizes (20 physician-presented vs. 4 patient/individual-user-presented videos), these between-group comparisons were considered exploratory. Associations between video characteristics, GQS, mDISCERN, and total content scores were assessed using Spearman rank correlation coefficient. Comment-related analyses included only videos for which comment counts were available. A two-sided $P < 0.05$ was considered statistically significant. Statistical analyses were performed using Python version 3.13.5 (Python Software Foundation, Wilmington, DE, USA), with SciPy version 1.17.0.

Results

Search Results and Video Characteristics

The initial six YouTube searches yielded 55 results before duplicate removal. After duplicates were identified and removed using unique YouTube video IDs, 22 unique videos remained. Six videos were excluded during eligibility screening, leaving 16 eligible videos from the initial search strategy.

The three additional anti-TNF searches yielded 91 search results in total (31, 26, and 34, respectively), corresponding to 52 unique videos after duplicate removal. Most were non-Turkish-language videos. Nine additional Turkish-language candidate videos were identified, of which eight met the eligibility criteria. Consequently, a total of 24 Turkish-language YouTube videos were included in the final analysis.

Of the included videos, 20 (83.3%) were physician-presented and 4 (16.7%) were presented by patients or individual users.

The median video duration was 1.93 minutes (IQR, 1.38–6.51; range, 0.90–67.13 minutes). The median number of views was 8,650 (IQR, 5,850–17,000; range, 15–100,000), and the median number of likes was 42 (IQR, 15.25–143.25; range, 0–442). Comment counts were available for 11 videos; among these, the median number of comments was 13 (IQR, 2.5–43; range, 0–289) (Table 1).

Content Assessment

The median total content score was 5 (range, 1–10), and none of the videos addressed all 11 predefined content domains. Indications/rheumatic diseases in which biologic therapies are

used was the most frequently addressed domain (19/24, 79.2%), whereas pregnancy/breastfeeding was the least frequently addressed (1/24, 4.2%). Detailed coverage of the individual content domains is presented in Table 2.

Table 1. General Characteristics, Quality, Reliability, and Usefulness of the Included YouTube Videos (N = 24)

Variable	Value
Presenter type	
Physician-presented	20 (83.3%)
Patient/individual-user-presented	4 (16.7%)
Video duration (minutes)	1.93 (1.38–6.51; 0.90–67.13)
Views	8,650 (5,850–17,000; 15–100,000)
Likes	42 (15.25–143.25; 0–442)
Comments (n = 11)	13 (2.5–43; 0–289)
Total content score (range 0–11)	5 (1–10)
mDISCERN score (range 0–5)	2 (1–4)
Score 1	2 (8.3%)
Score 2	12 (50.0%)
Score 3	8 (33.3%)
Score 4	2 (8.3%)
Score 5	0 (0.0%)
GQS (range 1–5)	3 (1–5)
Low quality (GQS 1–2)	1 (4.2%)
Moderate quality (GQS 3)	15 (62.5%)
High quality (GQS 4–5)	8 (33.3%)
Usefulness classification	
Useful medical information	18 (75.0%)
Misleading medical information	2 (8.3%)
Useful patient experience/opinion	3 (12.5%)
Misleading patient experience/opinion	1 (4.2%)

Data are presented as median (IQR; range) or n (%), as applicable. IQR is not reported for total content score, mDISCERN score, or GQS, for which only the median (range) was given. Comment counts were available for 11 of the 24 videos. Abbreviations: IQR: interquartile range, mDISCERN: modified DISCERN, GQS: Global Quality Score.

Table 2. Coverage of the 11-Item Biologic Therapy Content Checklist Across the Included Videos (N = 24)

Content domain	Videos addressing domain, n (%)
1. Definition/general characteristics of biologic therapy	9 (37.5%)
2. Mechanism of action	12 (50.0%)
3. Indications/rheumatic diseases in which biologic therapies are used	19 (79.2%)
4. Criteria for treatment initiation/patient selection	14 (58.3%)
5. Route and/or frequency of administration	8 (33.3%)
6. Pre-treatment infection screening/evaluation and/or vaccination recommendations	14 (58.3%)
7. Risk of infection	17 (70.8%)
8. Other adverse events	9 (37.5%)
9. Contraindications, precautions, and/or concomitant use with other rheumatologic treatments	10 (41.7%)
10. Monitoring/follow-up during treatment	10 (41.7%)
11. Pregnancy/breastfeeding	1 (4.2%)

Data are presented as n (%). The total content score (range, 0–11) was calculated as the sum of the 11 binary domain scores (present = 1, absent = 0). No video addressed all 11 domains.

Data are presented as n (%). The total content score (range, 0–11) was calculated as the sum of the 11 binary domain scores (present = 1, absent = 0). No video addressed all 11 domains.

Video Quality and Reliability

The median mDISCERN score was 2 (range, 1–4), and no video achieved the maximum score of 5. The median GQS was 3 (range, 1–5), indicating moderate overall quality. Detailed score distributions are presented in Table 1.

Usefulness Classification

According to the overall usefulness classification, 18 of the 24 videos (75.0%) were categorized as useful medical information, two (8.3%) as misleading medical information, three (12.5%) as useful patient experience/opinion, and one (4.2%) as misleading patient experience/opinion.

Overall, 21 videos (87.5%) were therefore classified as providing useful medical information or useful patient experiences/opinions, whereas three (12.5%) were classified as misleading.

Table 3. Correlations Among Video Characteristics and Quality, Reliability, and Content Measures; Comparison by Presenter Type; and Inter-Rater Agreement

Analysis	Physician-presented (n = 20), median (IQR)	Patient/individual-user-presented (n = 4), median (IQR)	Statistic	P-value
Correlations (Spearman's ρ)				
GQS vs. mDISCERN score	—	—	$\rho = 0.452$	$P = 0.026$
GQS vs. total content score	—	—	$\rho = 0.436$	$P = 0.033$
mDISCERN score vs. total content score	—	—	$\rho = 0.375$	$P = 0.071$
Video duration vs. total content score	—	—	$\rho = 0.801$	$P < 0.001$
Video duration vs. GQS	—	—	$\rho = 0.474$	$P = 0.019$
Video duration vs. mDISCERN score	—	—	$\rho = 0.237$	$P = 0.264$
Views vs. GQS	—	—	$\rho = -0.185$	$P = 0.387$
Views vs. mDISCERN score	—	—	$\rho = 0.120$	$P = 0.577$
Views vs. total content score	—	—	$\rho = -0.241$	$P = 0.257$
Likes vs. GQS	—	—	$\rho = 0.076$	$P = 0.724$
Likes vs. mDISCERN score	—	—	$\rho = 0.300$	$P = 0.155$
Likes vs. total content score	—	—	$\rho = 0.136$	$P = 0.527$
Comments (n = 11) vs. total content score	—	—	$\rho = -0.608$	$P = 0.047$
Comments (n = 11) vs. GQS	—	—	$\rho = -0.354$	$P = 0.286$
Comments (n = 11) vs. mDISCERN score	—	—	$\rho = -0.177$	$P = 0.603$
Presenter-type comparison				
Total content score	5.00 (3.00–6.25)	4.50 (3.50–5.50)	U = 46.5	$P = 0.637$
mDISCERN score	2.00 (2.00–3.00)	2.00 (1.75–2.25)	U = 53.0	$P = 0.290$
GQS	3.00 (3.00–4.00)	3.50 (2.50–4.00)	U = 41.0	$P = 0.964$
Number of views	8,650 (5,850–13,500)	31,400 (6,003.25–66,250)	U = 29.0	$P = 0.416$
Number of likes	33 (15.25–73)	215 (154.75–280)	U = 20.5	$P = 0.141$
Video duration, min	1.81 (1.29–5.75)	5.32 (2.37–15.04)	U = 23.0	$P = 0.201$
Inter-rater agreement				
11-item content checklist, % observed agreement	—	—	87.1%	—
11-item content checklist, pooled Cohen's κ	—	—	$\kappa = 0.735$ (95% CI 0.655–0.809)	—
Individual mDISCERN items, % observed agreement	—	—	80.8%	—
Individual mDISCERN items, pooled Cohen's κ	—	—	$\kappa = 0.618$ (95% CI 0.487–0.746)	—
Total mDISCERN score, ICC(2,1)	—	—	ICC = 0.525 (95% CI -0.316–0.868)	—
GQS, quadratic-weighted Cohen's κ	—	—	$\kappa = 0.412$ (95% CI 0.149–0.606)	—
Usefulness classification, Cohen's κ	—	—	$\kappa = 0.679$ (95% CI 0.371–1.000)	—

Data are Spearman rank correlation coefficients (ρ), Mann–Whitney U statistics, Cohen's κ , or the intraclass correlation coefficient (ICC), with exact P-values where applicable. Presenter-type data are presented as median (IQR), and these comparisons were considered exploratory because of the small and markedly unequal group sizes. Comment-related analyses included only videos for which comment counts were available (n = 11). For the 11-item content checklist and five individual mDISCERN items, pooled Cohen's κ was calculated by combining item-level ratings across all videos; 95% CIs were obtained using cluster bootstrap resampling at the video level. The ICC for the total mDISCERN score was based on a two-way random-effects, single-measure, absolute-agreement model [ICC(2,1)]. GQS agreement was assessed using quadratic-weighted Cohen's κ . Abbreviations: CI: confidence interval, GQS: Global Quality Score, mDISCERN: modified DISCERN, ICC: intraclass correlation coefficient, IQR: interquartile range. $P < 0.05$ was considered statistically significant.

Inter-Rater Agreement

For the 11-item content assessment, the overall observed agreement between the two reviewers was 87.1%, with a pooled Cohen's κ of 0.735 (95% CI 0.655–0.809). For the five individual mDISCERN items, the overall observed agreement was 80.8%, with a pooled Cohen's κ of 0.618 (95% CI 0.487–0.746).

The ICC for the total mDISCERN score was 0.525 (95% CI -0.316–0.868). Agreement for GQS was moderate (quadratic-weighted Cohen's $\kappa = 0.412$ [95% CI 0.149–0.606]), whereas agreement for usefulness classification was substantial (Cohen's $\kappa = 0.679$ [95% CI 0.371–1.000]). All disagreements were subsequently resolved by consensus, and consensus ratings were used for the final analyses.

Associations Between Quality, Reliability, and Content

GQS showed a moderate positive correlation with both the mDISCERN score (Spearman's $\rho = 0.452$, $P = 0.026$) and the total 11-item content score ($\rho = 0.436$, $P = 0.033$).

A positive correlation was also observed between the mDISCERN score and the total content score; however, this association did not reach statistical significance ($\rho = 0.375$, $P = 0.071$).

Video duration showed a strong positive correlation with the total content score ($\rho = 0.801$, $P < 0.001$) and a moderate positive correlation with GQS ($\rho = 0.474$, $P = 0.019$). In contrast, video duration was not significantly correlated with the mDISCERN score ($\rho = 0.237$, $P = 0.264$).

Associations With Video Popularity and Engagement

The number of views was not significantly correlated with GQS ($\rho = -0.185$, $P = 0.387$), mDISCERN score ($\rho = 0.120$, $P = 0.577$), or total content score ($\rho = -0.241$, $P = 0.257$).

Similarly, the number of likes was not significantly correlated with GQS ($\rho = 0.076$, $P = 0.724$), mDISCERN score ($\rho = 0.300$, $P = 0.155$), or total content score ($\rho = 0.136$, $P = 0.527$). Among videos for which comment counts were available (n = 11), the number of comments was negatively correlated with the total content score ($\rho = -0.608$, $P = 0.047$), whereas correlations with GQS ($\rho = -0.354$, $P = 0.286$) and mDISCERN score ($\rho = -0.177$, $P = 0.603$) were not statistically significant. Given the limited number of videos with available comment data, this finding was considered exploratory.

Comparison According to Presenter Type

When videos were compared according to presenter type, no statistically significant differences were detected between physician-presented (n = 20) and patient/individual-user-presented videos in terms of total content score ($P = 0.637$), mDISCERN score ($P = 0.290$), GQS ($P = 0.964$), number of views ($P = 0.416$), number of likes ($P = 0.141$), or video duration ($P = 0.201$).

Given the small and markedly unequal group sizes, these between-group comparisons were considered exploratory and should be interpreted cautiously. These results are summarized in Table 3.

Discussion

In this cross-sectional study of 24 Turkish-language YouTube videos on biologic therapies in rheumatic diseases, we found that most videos were physician-presented and were generally classified as useful, yet overall quality was moderate, reliability was limited, and information content was frequently incomplete. To our knowledge, this is the first study to evaluate Turkish-language YouTube content specifically addressing

biologic therapies across the spectrum of rheumatic diseases, and the first in this specific area to systematically quantify content completeness using a dedicated multi-domain checklist rather than relying on quality and reliability scores alone.

The median GQS of 3 indicates moderate overall video quality, consistent with the variable and generally moderate quality reported in previous YouTube-based evaluations of rheumatology-related content [2, 3, 4, 5, 6, 16]. Importantly, although most videos in our sample were physician-presented, overall quality was not consistently high, suggesting that physician presentation alone does not ensure high-quality educational content.

The median mDISCERN score of 2 (of a possible 5), with no video achieving the maximum score, indicates limited reliability and is broadly consistent with the generally low mDISCERN scores reported in previous YouTube-based evaluations [2, 3, 4, 6, 17], although higher scores have also been reported [16]. Because mDISCERN specifically captures the transparency of information sources and the balance of the material presented rather than independently verifying the clinical accuracy or completeness of all medical statements, a video may be classified as useful while still receiving a low mDISCERN score. This distinction may explain why most videos in our sample were classified as useful despite generally low mDISCERN scores.

In our sample, 87.5% of videos were classified as useful, a finding broadly consistent with previous YouTube-based studies in related areas [4, 16, 17]. However, usefulness classification should not be interpreted as evidence of comprehensive content or high reliability, as demonstrated by the limited mDISCERN scores and uneven coverage of the predefined content domains in our study.

Most videos in our sample were physician-presented (83.3%), consistent with previous rheumatology-focused YouTube studies [3, 4, 5, 6, 16]. However, we found no statistically significant differences between physician-presented and patient/individual-user-presented videos in the evaluated measures. Given the small and markedly unequal group sizes (20 vs. 4), these exploratory comparisons are inconclusive and should not be interpreted as evidence of equivalence or as indicating that presenter type does not influence video quality or reliability.

Neither the number of views nor the number of likes was significantly associated with GQS, mDISCERN, or total content score in our sample. A similar absence of association between popularity metrics and video quality has been reported by Kocyigit and Akaltun [5] and by Zengin and Onder [3], suggesting that viewer engagement, as measured by views and likes, does not reliably signal educational quality on this topic. Other studies, however, have described weak or even paradoxical relationships between engagement metrics and quality indicators [2, 16, 18], underscoring that popularity should not be used as a proxy for the accuracy or completeness of medical information on YouTube. Among the subset of videos with available comment data ($n = 11$), comment count was negatively correlated with total content score; however, given the small sample and the multiple correlation analyses performed, this finding should be considered exploratory and interpreted cautiously.

Video duration was strongly correlated with the total content score and moderately correlated with GQS, but was not significantly correlated with mDISCERN. This pattern indicates that longer videos tended to cover more of the predefined clinical domains and were judged somewhat higher in overall quality, but were not thereby more transparent about their sources or more balanced in their presentation. This association may partly reflect the greater opportunity for longer videos to address a larger number of predefined content domains and should not be interpreted as evidence that increasing video duration itself improves educational quality. In other words, duration appears to track content comprehensiveness rather than information reliability, and the two should not be conflated: a longer video is not necessarily a more reliable one, even though it is more likely to be a more complete one.

A distinguishing feature of the present study is the use of an 11-item checklist to characterize which specific clinical domains were addressed, rather than relying solely on aggregate quality and reliability scores. This approach showed that no video addressed all 11 domains and that coverage was highly uneven: indications for biologic therapy and infection risk were the most frequently discussed domains, whereas route/frequency of administration, monitoring during treatment, and, most strikingly, pregnancy and breastfeeding were addressed only infrequently. Pregnancy/breastfeeding was discussed in only a single video (4.2%). Reproductive health is a clinically important consideration in the use of antirheumatic therapies, as reflected in contemporary rheumatology guidance on reproduction, pregnancy, and lactation [19, 20]. However, these guidelines are cited here to establish the clinical relevance of this topic rather than to imply that every patient-facing video should reproduce the scope or content of professional guidance. This gap is consistent with the pregnancy-specific focus of Yagiz et al. and Yerlanova et al., both of whom highlighted substantial deficiencies in the availability and quality of pregnancy-related rheumatology content on YouTube [4, 18]. These findings suggest that Turkish-language YouTube content on biologic therapies, while broadly informative regarding indications and infection-related risk, leaves clinically relevant gaps that could be particularly consequential for patients of reproductive age.

Inter-rater agreement varied across the assessment measures, with higher agreement estimates for the content checklist and individual mDISCERN items and more moderate agreement for the total mDISCERN score and GQS.

Overall, these findings should be interpreted in light of the small final sample size ($n = 24$), which limits statistical power, particularly for subgroup comparisons such as that between physician- and patient/individual-user-presented videos, and precludes firm causal or generalizable conclusions. As in other single-language, single-platform assessments of this kind [1, 2, 3, 4, 6, 16, 17, 18], our results describe a specific and time-limited snapshot of the Turkish-language YouTube landscape for this topic rather than a definitive characterization of it, and should be interpreted accordingly.

This study has several limitations. First, the study-specific content checklist was not a validated instrument, and all domains contributed equally to the total content score despite potential differences in their clinical importance. Therefore, the

total content score should be interpreted as a measure of the breadth of predefined topic coverage rather than as a weighted measure of clinical completeness. Second, the final sample comprised only 24 videos, and the patient/individual-user-presented subgroup consisted of only 4 videos; the resulting limited statistical power means that non-significant comparisons, particularly by presenter type, cannot be interpreted as evidence of no difference. Third, YouTube content, along with view counts, like counts, and comment counts, is inherently dynamic, and videos may be added, removed, edited, or re-ranked over time; our findings therefore reflect a single cross-sectional snapshot at the time of the search rather than a stable or permanent state of the platform. Fourth, search results on YouTube may vary depending on the search date, the device or browser used, region, and the platform's proprietary ranking algorithm, even when steps are taken to minimize personalization, such as searching in a private browsing mode without an active account; our results may not be fully reproducible with a different search date or configuration. Fifth, the study was restricted to Turkish-language videos identified using a predefined set of Turkish search terms; videos using alternative terminology, dialectal variation, or content behind different keywords may have been missed, and the findings cannot be generalized to other languages or platforms. Sixth, comment counts were available for only 11 of the 24 videos, limiting the statistical power and generalizability of comment-related analyses, which should accordingly be regarded as exploratory. Finally, although the content checklist, mDISCERN, and GQS were applied independently by two rheumatologists with disagreements resolved by consensus, quality and reliability assessment inherently retains a subjective component, and some degree of rater-dependent variability, particularly given the variation in inter-rater agreement across measures, cannot be entirely excluded.

Conclusion

In this study of Turkish-language YouTube videos on biologic therapies in rheumatic diseases, most videos were physician-presented and were generally classified as useful, but overall quality was moderate, reliability was limited, and information content was frequently incomplete, most notably with respect to pregnancy and breastfeeding. Video duration was strongly associated with content comprehensiveness but not with reliability, and neither views nor likes were associated with quality or reliability, indicating that popularity is not a reliable proxy for educational value on this topic. Because the sample was small and the patient/individual-user subgroup limited to 4 videos, these findings should be regarded as preliminary and hypothesis-generating rather than conclusive, and should not be extrapolated causally or to other languages, platforms, or time periods. Notably, although most videos were physician-presented, physician presentation alone did not ensure high reliability or broad coverage of the predefined content domains. These findings highlight the need for Turkish-language educational content on biologic therapies that provides broader coverage of clinically relevant topics, particularly reproductive health, administration, and treatment monitoring.

References

- Kocyigit BF, Akyol A. YouTube as a source of information on COVID-19 vaccination in rheumatic diseases. *Rheumatol Int.* 2021;41(12):2109-15. <https://doi.org/10.1007/s00296-021-05010-2>.
- Zengin O, Onder ME. Educational quality of YouTube videos on musculoskeletal ultrasound. *Clin Rheumatol.* 2021;40(10):4243-51. <https://doi.org/10.1007/s10067-021-05793-6>.
- Zengin O, Onder ME. YouTube for information about side effects of biologic therapy: a social media analysis. *Int J Rheum Dis.* 2020;23(12):1645-50. <https://doi.org/10.1111/1756-185X.14003>.
- Yagiz B, Coskun BN, Yunus Halil E, Dalkilic E, Pehlivan Y. The efficacy and reliability of English YouTube videos as a source of information for pregnant rheumatoid arthritis patients. *Clin Rheumatol.* 2023;42(12):3311-20. <https://doi.org/10.1007/s10067-023-06780-9>.
- Kocyigit BF, Akaltun MS. Does YouTube provide high quality information? Assessment of secukinumab videos. *Rheumatol Int.* 2019;39(7):1263-8. <https://doi.org/10.1007/s00296-019-04322-8>.
- Koçak Ulucaköy R, Kayacan Erdoğan E. A review of YouTube videos in the Turkish language about ankylosing spondylitis. *Türkiye Klinikleri J Phys Med Rehabil Sci.* 2025;28(1):63-9. <https://doi.org/10.31609/jpmrs.2024-106018>.
- Drozdz B, Couvillon E, Suarez A. Medical YouTube videos and methods of evaluation: literature review. *JMIR Med Educ.* 2018;4(1):e3. <https://doi.org/10.2196/mededu.8527>.
- Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. *Health Informatics J.* 2015;21(3):173-94. <https://doi.org/10.1177/1460458213512220>.
- Pamukcu M, Izci Duran T. Are YouTube videos enough to learn anakinra self-injection? *Rheumatol Int.* 2021;41(12):2125-31. <https://doi.org/10.1007/s00296-020-04999-w>.
- Tolu S, Yurdakul OV, Basaran B, Rezvani A. English-language videos on YouTube as a source of information on self-administer subcutaneous anti-tumour necrosis factor agent injections. *Rheumatol Int.* 2018;38(7):1285-92. <https://doi.org/10.1007/s00296-018-4047-8>.
- Elangovan S, Kwan YH, Fong W. The usefulness and validity of English-language videos on YouTube as an educational resource for spondyloarthritis. *Clin Rheumatol.* 2021;40(4):1567-73. <https://doi.org/10.1007/s10067-020-05377-w>.
- Ng CH, Lim GRS, Fong W. Quality of English-language videos on YouTube as a source of information on systemic lupus erythematosus. *Int J Rheum Dis.* 2020;23(12):1636-44. <https://doi.org/10.1111/1756-185X.13852>.
- Delli K, Livas C, Vissink A, Spijkervet FKL. Is YouTube useful as a source of information for Sjögren's syndrome? *Oral Dis.* 2016;22(3):196-201. <https://doi.org/10.1111/odi.12404>.
- Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. *Am J Gastroenterol.* 2007;102(9):2070-7. <https://doi.org/10.1111/j.1572-0241.2007.01325.x>.
- Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health.* 1999;53(2):105-11. <https://doi.org/10.1136/jech.53.2.105>.
- Coşkun BN, Yagiz B, Giounous Chalil E, Dalkılıç E, Pehlivan Y. The usefulness and reliability of English-language YouTube videos as a source of knowledge for patients with familial Mediterranean fever. *PeerJ.* 2024;12:e16857. <https://doi.org/10.7717/peerj.16857>.
- Seyyar M, Sahin E. Türkçe YouTube videolarında immünoterapi: kalite, güvenilirlik ve bilgi içeriğinin değerlendirilmesi. *Acta Med Nicomedia.* 2026;9:1-7. <https://doi.org/10.53446/actamednicomedia.1850758>.
- Yerlanova D, Zimba O, Khojakulova U, Yessirkepov M, Kocyigit BF. Analysis of YouTube videos on rheumatic diseases and pregnancy: a cross-sectional study. *Rheumatol Int.* 2026;46:68. <https://doi.org/10.1007/s00296-026-06102-7>.
- Rüegg L, Pluma A, Hamroun S, Cecchi I, Perez-Garcia LF, Anderson PO, et al. EULAR recommendations for use of antirheumatic drugs in reproduction, pregnancy, and lactation: 2024 update. *Ann Rheum Dis.* 2025;84(6):910-26. <https://doi.org/10.1016/j.ard.2025.02.023>.
- Sammaritano LR, Bermas BL, Chakravarty EE, Chambers C, Clowse MEB, Lockshin MD, et al. 2020 American College of Rheumatology guideline for the management of reproductive health in rheumatic and musculoskeletal diseases. *Arthritis Care Res (Hoboken).* 2020;72(4):461-88. <https://doi.org/10.1002/acr.24130>.

Disclaimer/Publisher's Note: The statements, opinions, and data presented in publications in the Journal of Surgery and Medicine (JOSAM) are exclusively those of the individual author(s) and contributor(s) and do not necessarily reflect the views of JOSAM, the publisher, or the editor(s). JOSAM, the publisher, and the editor(s) disclaim any liability for any harm to individuals or damage to property that may arise from implementing any ideas, methods, instructions, or products referenced within the content. Authors are responsible for all content in their article(s), including the accuracy of facts, statements, and citations. Authors are responsible for obtaining permission from the previous publisher or copyright holder if re-using any part of a paper (e.g., figures) published elsewhere. The publisher, editors, and their respective employees are not responsible or liable for the use of any potentially inaccurate or misleading data, opinions, or information contained within the articles on the journal's website.