

Research Article

Quality Variation of Osteogenesis Imperfecta Videos on YouTube: A Viewership-Based Evaluation

 Süleyman Kozlu,  Enes Alptekin Canlı

Department of Orthopedics and Traumatology, Kutahya Health Sciences University, Kutahya, Türkiye

Abstract

Objectives: To evaluate the educational quality, reliability, and timeliness of Osteogenesis Imperfecta (OI) videos on YouTube, examine relationships with popularity indicators (views, likes), and analyze quality differences across source types.

Methods: In May 2024, a standard YouTube search was conducted for "Osteogenesis Imperfecta" and "Brittle Bone Disease." The first 200 videos sorted by relevance were screened, and 27 eligible English videos were analyzed. Reliability was scored using JAMA Benchmark Criteria (1–4) and educational quality via Global Quality Score (GQS, 1–5). Sources were classified (academic, patient narrative, documentary) and compared using Kruskal–Wallis and Dunn tests.

Results: The overall mean JAMA and GQS scores were 2.27 and 3.02. Patient stories garnered the highest engagement but the lowest quality scores (JAMA 2.10; GQS 2.60). Academic videos were significantly superior to patient stories in both JAMA ($p=0.027$) and GQS ($p=0.013$). While views and likes strongly correlated ($r=0.93$), popularity indicators showed a weak, insignificant negative correlation with quality scores ($r= -0.14$ to -0.20).

Conclusion: YouTube OI videos offer moderate educational quality. Although academic content is scientifically superior, viewer popularity is heavily concentrated in lower-quality patient narratives. This disparity highlights an urgent need for institutions to produce more visible, validated content.

Keywords: GQS, JAMA, Osteogenesis imperfecta, Patient education, YouTube

Cite This Article: Kozlu S, Alptekin Canlı E. Quality Variation of Osteogenesis Imperfecta Videos on YouTube: A Viewership-Based Evaluation. EJMI 2026;10(1):60–64.

Osteogenesis imperfecta (OI) is a hereditary connective tissue disease caused by mutations in the COL1A1 and COL1A2 genes involved in type I collagen synthesis.^[1,2] These mutations disrupt the structural integrity of the bone matrix, leading to increased fragility, short stature, deformities, scoliosis, and joint hypermobility.^[3,4] The prevalence of OI ranges from approximately 1:15,000–1:20,000 live births.^[1,5]

This clinically heterogeneous disease is defined in forms ranging in severity from types I to IV according to the Silience classification. The spectrum has expanded with the subsequent subtypes of types V–VIII.^[6] OI can affect not only the skeletal system but also multiple organ systems, such as dentinogenesis imperfecta, hearing loss, blue sclera, and cardiac anomalies. Management of the disease requires a multidisciplinary approach; Orthopedic surgery, physiotherapy, bisphosphonate therapy, psychosocial sup-

Address for correspondence: Süleyman Kozlu, MD. Department of Orthopedics and Traumatology, Kutahya Health Sciences University, Kutahya, Türkiye

Phone: +90 542 632 22 92 **E-mail:** suleyman.kozlu@ksbu.edu.tr

Submitted Date: February 02, 2026 **Accepted Date:** March 08, 2026

©Copyright 2025 by Eurasian Journal of Medicine and Investigation - Available online at www.ejmi.org

OPEN ACCESS This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.



port, and genetic counseling are key components.^[7] However, the rarity of OI leads to low public awareness and lack of access to information.

Today, access to medical information is provided through digital media platforms rather than traditional sources. YouTube, with over 2.5 billion monthly users, is now the most popular place in the world to get visual information. While the platform's audiovisual content facilitates public understanding of diseases, lack of scientific oversight and disinformation pose serious risks.^[8]

Therefore, the aim of our study was to evaluate the educational quality, reliability, timeliness, and relationship to popularity criteria of Osteogenesis Imperfecta videos on YouTube. We also aimed to reveal quality differences between different source types (academic, professional, and patient experience).

Methods

Study Design

This study was designed as an observational, cross-sectional, and descriptive analysis. The study aimed to evaluate the educational quality, reliability, timeliness, and popularity indicators of videos related to Osteogenesis Imperfecta (OI) on the YouTube platform. Ethics committee approval was not required because the data were obtained from a publicly available platform. However, the study design complies with the principles of the Declaration of Helsinki.^[9]

Search Strategy

A YouTube search was conducted in May 2024 using standard browser mode without logging and with browsing history deleted. The keywords were "Osteogenesis Imperfecta" and "Brittle Bone Disease." The option to sort was set to "relevance." The first 200 videos obtained from the search were pre-screened, and those that met the inclusion criteria were included in the study.^[3,10]

Inclusion Criteria

- English language content
- Focus directly on the topic of Osteogenesis Imperfecta
- Have at least 2000 views
- Have sufficient visual and audio quality
- Have a clearly identified video source (e.g., academic institution, physician, professional organization).^[8]

Exclusion Criteria

- Advertisement, product promotion, or entertainment content

- Short-duration (<1 min) or off-topic videos
- Repetitive or copied content from other platforms
- Videos with clearly questionable scientific accuracy or conspiracy themes.^[8]

Video Classification

Videos were divided into four main categories based on content type:^[5]

- Academic Institutions
- Patient Stories
- Health Information Sites
- Educational Videos

Evaluation Scales

The most commonly used criteria for evaluating health videos are the JAMA Benchmark Criteria and the Global Quality Score (GQS). These two scales are considered reliable tools for measuring both the scientific credibility and educational quality of videos.

a) JAMA Benchmark Criteria[11] The credibility of videos was scored according to four parameters: • Author/producer identification, • Citation, • Date updated, • Declaration of financial interest. One point was given for each parameter, and the total score ranged from 1 to 4.

b) Global Quality Score (GQS)[12] A 5-point subjective scale was used to measure the overall educational value of the video: 1 = poor, 2 = limited, 3 = fair, 4 = good, 5 = excellent quality. This method is frequently used to evaluate the educational integrity of health-related videos.

Data Collection Process

The following parameters were recorded for each video: • Title, upload year, duration (seconds), • Number of views, number of likes, number of comments, • Source type, • JAMA and GQS scores. Scoring was performed by two independent raters (orthopedic and traumatology specialists). We used Cohen's kappa coefficient (κ) to measure inter-rater consistency. Values between 0.81 and 1.00 were seen as "high agreement".^[13]

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics v.26.0 (IBM Corp., Armonk, NY, USA). Mean \pm standard deviation was calculated for continuous variables, and percentages (%) for categorical variables. Normal distribution of the data was assessed using the Shapiro–Wilk test. Differences between groups were analyzed using the Kruskal–Wallis test, and $p < 0.05$ was considered significant.^[14]

Results

A total of 27 videos were evaluated. The videos had an average JAMA score of 2.27 and a GQS score of 3.02, which means the content quality was average. When we looked at the different types of sources, we saw that patient story videos (type 2) had the most views and likes. Type 3 health information websites and type 4 educational videos had better content, as shown by their higher JAMA and GQS scores. Videos from academic institutions (type 1) didn't last long and didn't get as much attention. The results of the evaluation of videos by source type are presented in Table 1.

The correlation analysis found a very strong positive correlation ($r=0.93$) between the number of views and the number of likes. This result suggests that videos with higher views also garner more likes, which is an expected result. Additionally, a moderate-to-high positive correlation ($r=0.61$) was found between the JAMA score and the GQS score, indicating that video quality improves in parallel with both academic and general evaluation criteria. Conversely, weak negative correlations were found between the JAMA score and likes ($r=0.20$), the GQS score and views ($r=-0.19$), and the JAMA score and views ($r=-0.14$). Correlation according to variables is shown in Figure 1.

This suggests that scientifically high-quality videos do not receive the same level of attention in terms of views or likes; viewer interest may be more related to factors such as presentation style, emotional impact, or storytelling than to academic content. The highest and lowest correlation variables and their interpretations by group are shown in Tables 2 and 3.

Views and likes are the variables with the strongest positive correlation ($r=0.93$). JAMA and GQS scores are strongly positively correlated with each other ($r=0.61$). There is a slight negative trend between quality scores (JAMA, GQS) and the number of views or likes. This suggests that videos with higher scientific quality may attract less attention in terms of popularity. These results suggest that there is no direct relationship between popularity metrics (views, likes) and scientific quality indicators (JAMA, GQS), with

user interest being more directed towards storytelling and emotional content.

The Kruskal–Wallis test revealed a statistically significant difference between the groups in terms of JAMA scores ($\chi^2=7.41$, $p=0.027$). Similarly, a significant difference was observed in terms of GQS scores ($\chi^2=8.72$, $p=0.013$). According to the multiple comparison analysis (Dunn post-hoc test):

- Both JAMA and GQS scores of academic videos were significantly higher than those of patient stories ($p=0.021$ and $p=0.009$).
- No significant difference was observed between academic and professional (health information or education) videos ($p=0.312$).

However, viewer interest and engagement were generally directed towards patient stories, which are generally low in scientific accuracy.

Discussion

This study is one of the few systematic studies examining the educational quality, reliability, timeliness, and relationship between popularity indicators and Osteogenesis Imperfecta (OI) videos on YouTube. Today, digital media, particularly YouTube, has become one of the most common

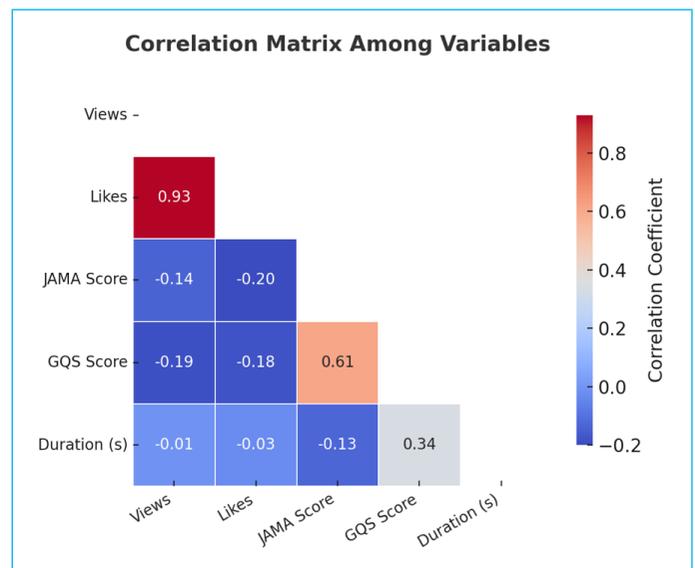


Figure 1. Variable correlation.

Table 1. Comparison by source type

| Source type | n | Mean views | Mean likes | JAMA score | GQS score | Mean duration (s) |
|--------------------------|----|------------|------------|------------|-----------|-------------------|
| Academic institutions | 3 | 24,533 | 259 | 2.00 | 3.00 | 150 |
| Patient stories | 10 | 1,555,510 | 32,390 | 2.10 | 2.60 | 356 |
| Health information sites | 9 | 45,022 | 518 | 2.56 | 3.33 | 653 |
| Educational videos | 5 | 20,580 | 236 | 2.40 | 2.80 | 455 |
| Overall (total) | 27 | 445,858 | 9,277 | 2.27±0.33 | 3.02±0.58 | 419 |

Table 2. Highest positive correlations among variables

| Variable pair | Correlation (r) | Interpretation |
|------------------------|-----------------|---|
| Views - likes | 0.93 | There is a very strong positive correlation between the number of views and the number of likes on videos. |
| JAMA score - GQS score | 0.61 | There is a moderate to strong positive relationship, which means that academic and general quality scores go up at the same time. |
| GQS score - duration | 0.34 | There is a weak positive correlation between longer videos and higher GQS scores. |

Table 3. Highest negative correlations among variables

| Variable pair | Correlation (r) | Interpretation |
|--------------------|-----------------|---|
| JAMA score - Likes | -0.20 | There is a small negative correlation: videos that are better academically tend to get fewer likes. |
| GQS score - Views | -0.19 | Weak negative correlation: more views don't mean higher general quality scores. |
| JAMA score-Views | -0.14 | Weak negative correlation: scientifically rigorous videos draw fewer viewers. |

sources for accessing health information. With over 2.5 billion monthly users, this platform has great potential for public health awareness but carries a serious risk of disinformation due to a lack of scientific oversight. The fact that patients, especially those with rare diseases—such as OI—turn to digital platforms for accurate information further enhances the importance of online content quality.

The findings indicate that OI videos generally offer a moderate level of reliability (mean JAMA score=2.27) and moderate educational quality (mean GQS=3.02). This result is consistent with previously reported findings for orthopedic and musculoskeletal content on YouTube. Koller et al. reported that the majority of hip arthritis videos provided insufficient information regarding diagnosis and treatment; Singh et al. reported that over 70% of rheumatoid arthritis videos contained inaccurate or incomplete information.^[5,15] Consistent with these studies, our study also found that videos with high scientific accuracy had low viewing rates, while more popular videos were insufficiently educational. A similar correlation was reported in a study by Osman et al.^[8] between compliance with JAMA criteria and viewers' trust in the content.

Comparisons by source type revealed that academic content had significantly higher JAMA and GQS scores than patient story videos ($p=0.021$ and $p=0.009$). This finding is consistent with the study by Oztermeli and Karahan^[16] on developmental hip dysplasia. However, no significant difference was observed between academic and professional (health information or education-based) videos ($p=0.312$). This demonstrates that professional sources can produce content as reliable as academic institutions.

In our study, patient story videos received the lowest quality scores (JAMA=2.10; GQS=2.60), despite having the highest number of views and likes. Correlation analyses showed

that the number of views and likes was weakly negatively correlated with the JAMA and GQS scores ($r=-0.14$ to -0.20 , $p>0.05$). This suggests that popularity is related to emotional narrative, personal story, and viewer empathy rather than scientific accuracy. So, health organizations and schools should put as much emphasis on narrative techniques that keep people interested as they do on scientific accuracy. Visual explanations, captioning options, the use of plain language, and case-based examples can ensure accurate information is conveyed while maintaining public interest.^[17]

Due to its rarity and lifelong multisystemic effects, OI creates a serious information gap for patients. The results of this study demonstrate that although patient experiences disseminated online enhance awareness, they are inadequate for transmitting scientific knowledge. Madathil et al.^[1] and Drozd et al.^[10] similarly emphasized that the majority of medical content on YouTube lacks academic integrity. In this context, it is important for academic centers, expert associations, and professional communities to become active content producers on digital platforms.

Furthermore, the use of “verified medical creator” labels on platforms like YouTube can increase information security. A key limitation of this study is the relatively small sample size ($n=27$) and the YouTube algorithm's ability to vary over time. Furthermore, the JAMA and GQS assessments include subjective observer-dependent criteria. Another limitation is that only videos in English were evaluated; content in other languages was excluded.

Conclusion

Osteogenesis Imperfecta videos on YouTube are generally of moderate educational quality but vary significantly in terms of source type and content up-to-dateness. Aca-

demical and professional content offers the most scientifically credible videos, while patient story videos garner more public attention. This highlights the disconnect between popularity and scientific accuracy. To raise public health awareness, it is crucial for academic institutions and professional associations to take an active role on digital platforms and produce content that balances scientific accuracy with accessibility.

Disclosures

Ethics Committee Approval: Ethics committee approval was not required because the data were obtained from a publicly available platform.

Conflicts of Interest: The authors declare that they have no competing interests.

Funding: The author declared that this study has received no financial support.

Use of AI for Writing Assistance: Generative AI tools (Gemini, Google) were used for language editing, translation assistance, and structural formatting of the manuscript. The authors reviewed and take full responsibility for the content.

Authors' contributions: Concept – SK; Design – SK, EAC; Supervision – SK; Materials – SK, EAC; Data Collection and/or Processing – SK, EAC; Analysis and/or Interpretation – SK; Literature Review – SK, EAC; Writing – SK, EAC; Critical Review – SK, EAC.

Peer-review: Externally peer reviewed.

References

1. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: A systematic review. *Health Informatics J* 2015;21(3):173–94. [\[CrossRef\]](#)
2. Anand S, Singh A, Singh V, Sharma S. Comprehensive assessment of the quality and reliability of the ten most-viewed YouTube videos on thoracoscopic lobectomy in children: a comparison from the available videos on a peer-reviewed platform. *Pediatr Surg Int* 2021;37(11):1627–32. [\[CrossRef\]](#)
3. Basch C, Hillyer G, Meleo-Erwin Z, Zagnit E. YouTube videos and informed decision-making about COVID-19 vaccination: Successive sampling study. *JMIR Public Health Surveill* 2021;7(5):e28352. [\[CrossRef\]](#)
4. Desai T, Shariff A, Dhingra V, Minhas D, Eure M, Kats M. Is content really king? An objective analysis of the public's response to medical videos on YouTube. *PLoS One* 2013;8(12):e82469. [\[CrossRef\]](#)
5. Koller U, Waldstein W, Schatz KD, Windhager R. YouTube provides irrelevant information for the diagnosis and treatment of hip arthritis. *Int Orthop* 2016;40(10):1995–2002. [\[CrossRef\]](#)
6. Van Dijk FS, Pals G, Van Rijn RR, Nikkels PG, Cobben JM. Classification of Osteogenesis Imperfecta revisited. *Eur J Med Genet* 2010;53(1):1–5. [\[CrossRef\]](#)
7. Bregou Bourgeois A, Aubry-Rozier B, Bonafé L, Laurent-Aplegate L, Pioletti DP, Zambelli PY. Osteogenesis imperfecta: from diagnosis and multidisciplinary treatment to future perspectives. *Swiss Med Wkly* 2016;146:w14322. [\[CrossRef\]](#)
8. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med Educ* 2022;22(1):382. [\[CrossRef\]](#)
9. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013;310(20):2191–4. [\[CrossRef\]](#)
10. Drozd B, Couvillon E, Suarez A. Medical YouTube Videos and methods of evaluation: Literature review. *JMIR Med Educ* 2018;4(1):e3. [\[CrossRef\]](#)
11. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the internet: caveat lector et viewer—let the reader and viewer beware. *JAMA* 1997;277(15):1244–5. [\[CrossRef\]](#)
12. 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. [\[CrossRef\]](#)
13. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276–82. [\[CrossRef\]](#)
14. Kim TK. T test as a parametric statistic. *Korean J Anesthesiol* 2015;68(6):540–6. [\[CrossRef\]](#)
15. Singh AG, Singh S, Singh PP. YouTube for information on rheumatoid arthritis—a wakeup call?. *J Rheumatol* 2012;39(5):899–903. [\[CrossRef\]](#)
16. Oztermeli A, Karahan N. Evaluation of YouTube video content about developmental dysplasia of the hip. *Cureus* 2020;12(8):e9557. [\[CrossRef\]](#)
17. Covington P, Adams J, Sargin E. Deep Neural Networks for YouTube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*. New York, NY, USA: Association for Computing Machinery; 2016. p. 191–8. [\[CrossRef\]](#)