British Journal of Hospital Medicine

Br. J. Hosp. Med. (Lond) **2026**; 87(7): 51256
https://doi.org/10.31083/BJHM51256

Article

Evaluation of the Educational Quality and Reliability of Cardiopulmonary Resuscitation Training Videos on YouTube™

Süleyman Nogay[1,*], Turgut Dolanbay[1], Bilgehan Demir[1], Bilal Kartal[1], Asiye Özkan[1]

[1]Department of Emergency Medicine, Faculty of Medicine, Malatya Turgut Ozal University, 44210 Malatya, Türkiye
*Correspondence: drsnogay@gmail.com (Süleyman Nogay)

Academic Editor: John Sandars
Submitted: 26 February 2026 Revised: 16 April 2026 Accepted: 29 May 2026 Published: 20 July 2026

Abstract

Aims/Background: Cardiac arrest is a leading cause of mortality worldwide, and early, high-quality cardiopulmonary resuscitation (CPR) is critical for survival. With increasing use of digital media, YouTube™ has become a common resource for CPR education. However, the educational quality and scientific reliability of these videos remain inconsistent. This study aimed to evaluate the educational quality, reliability, and technical adequacy of CPR training videos on YouTube™ using tools and an evaluation index. **Methods:** A total of 120 CPR-related YouTube™ videos were analyzed. Videos were assessed using four validated tools: Modified DISCERN Scale (MDS), Journal of the American Medical Association (JAMA) scoring system, Global Quality Scale (GQS), and a newly developed evaluation index [cardiopulmonary resuscitation-video evaluation index (CPR-VEI)]. Video duration, views, likes, and uploader type were recorded, and analyses used nonparametric tests and Spearman correlation. **Results:** Institutionally produced videos demonstrated significantly higher view counts ($p = 0.008$) as well as higher JAMA ($p < 0.001$), CPR-VEI ($p = 0.020$), and MDS ($p = 0.031$) scores compared with individually uploaded content. Intermediate-duration videos demonstrated significantly higher CPR-VEI, GQS, and MDS scores compared with short-duration videos, while long-duration videos showed higher numbers of likes. Correlation analysis revealed strong positive associations between CPR-VEI and GQS and MDS scores, and a moderate positive correlation with the JAMA score. **Conclusion:** CPR training videos on YouTube™ exhibit substantial variability in educational quality and reliability. Institutionally produced videos demonstrated higher quality and reliability, while intermediate-duration videos showed higher educational scores and long-duration videos demonstrated greater engagement. However, popularity metrics such as views and likes did not reliably reflect educational quality. These findings support the role of guideline-based, technology-enhanced learning as an essential component of contemporary CPR education.

Keywords: cardiopulmonary resuscitation; social media; health education; educational measurement; emergency medical services

1. Introduction

Cardiac arrest remains one of the leading causes of morbidity and mortality worldwide, with early and effective cardiopulmonary resuscitation (CPR) being the most critical determinant of survival [1]. Globally, survival rates for out-of-hospital cardiac arrest remain below 10%. Early initiation of basic life support interventions significantly improves these survival outcomes [2]. Ensuring that both healthcare professionals and the general public have easy access to accurate and up-to-date information about CPR plays a crucial role in enhancing the effectiveness of interventions and improving survival rates [2,3].

With the impact of the digital age, patterns of access to information have changed substantially. Video-sharing platforms, particularly YouTube™, have increasingly been used for health education and awareness purposes [3]. YouTube™ is an open-access platform that enables users to upload, view, comment on, and express approval of an unlimited range of video content. However, the lack of any formal scientific review or editorial oversight may lead to the dissemination of inaccurate, incomplete, or misleading information [4,5].

Recent studies have demonstrated that the quality and reliability of health-related videos on YouTube™ are highly variable [6,7]. A considerable proportion of CPR training videos have been reported to contain notable deficiencies in instructional content and to be only partially compliant with established resuscitation guidelines [4,8]. Furthermore, indicators of video popularity-such as the number of views and likes-have often been found not to correlate with scientific accuracy or educational value [5]. In the present study, video content was evaluated in accordance with current resuscitation principles, taking into account the framework outlined in the 2025 European Resuscitation Council (ERC) Guidelines.

Therefore, in light of these concerns, it is essential to evaluate the educational value of CPR videos published on YouTube™ using objective and standardized criteria. For this purpose, scoring systems such as the Modified DISCERN Scale (MDS), the Journal of the American Medical Association (JAMA) benchmark, and the Global Quality Scale (GQS) have been widely used to assess videos in terms of scientific accuracy, reliability, and presentation quality [9,10].

In addition to the commonly used scoring systems, we developed the cardiopulmonary resuscitation-video evaluation index (CPR-VEI) scoring system to enable a more objective, structured, and standardized assessment of CPR-related videos. The framework of this system was inspired by the structured skill-evaluation approach of the Objective Structured Assessment of Technical Skills (OSATS) scoring index.

The aim of this study was to evaluate the educational quality and reliability of CPR instructional videos available on the YouTube™ platform. By assessing their compliance with scientific standards, this evaluation is intended to contribute to the development of reliable educational materials in digital environments, thereby supporting public health and promoting effective resuscitation training.

2. Methods

2.1 Study Design and Setting

All data evaluated in this study were obtained from publicly available, open-access YouTube™ videos. No direct interaction with individuals occurred during the research process, no medical interventions were performed, and no identifiable personal data were used. Therefore, informed consent was not required, and institutional ethics committee approval was not necessary.

2.2 Video Selection and Evaluation Process

The video search was conducted on the YouTube™ platform on 10 October 2025. The keywords "cardiopulmonary resuscitation", "cardiac arrest resuscitation", "life support", and "emergency resuscitation" were entered into the search bar to identify relevant videos. To minimize the influence of user history on the recommendation algorithm, all searches were performed in incognito mode using a browser with cleared cache and cookies, without logging into any user account. The search was conducted from a single geographic location using default platform settings. Search results were not further filtered or personalized beyond the default platform settings. For each search term, the first 100 videos sorted by relevance were screened, reflecting typical user behavior on YouTube™. A total of 400 videos were initially identified. After removing duplicates and clearly irrelevant videos, 140 videos remained for detailed assessment. Subsequently, 20 videos were excluded based on predefined inclusion and exclusion criteria, resulting in 120 videos included in the final analysis. Although some of the included videos may have been uploaded prior to the publication of the 2025 European Resuscitation Council (ERC) Guidelines, all videos were evaluated according to the most up-to-date resuscitation recommendations available at the time of analysis to ensure a standardized and contemporary benchmark for assessment.

Inclusion criteria: Videos that demonstrate all steps of the CPR technique, content with sufficient visual quality (≥480p), videos containing real CPR procedures, cadaveric or simulation-based CPR footage, and content presented in English or in other languages with English subtitles.

Exclusion criteria: Advertisements, product promotions, or humorous content; videos omitting one or more steps of the CPR sequence; recordings with poor audio or visual quality; duplicate uploads from the same creator; and videos shorter than 60 seconds.

For each video, the following parameters were recorded: video title, duration, upload date, number of views and likes; source of upload [institutional (university, hospital, emergency medicine associations) or individual (physician, student, or general user)]; and country of origin (developed countries and developing countries).

Video duration was categorized as short (1–5 minutes), medium (5–10 minutes), and long (>10 minutes) based on commonly used thresholds in the literature and the distribution of video lengths observed in the study sample, as no universally accepted standard exists for duration classification in YouTube-based studies [11].

Five independent reviewers (three emergency medicine specialists and two emergency medicine residents) systematically evaluated all videos for adherence to CPR technique components, including chest compression depth and rate, hand position, ventilation effectiveness, chest recoil, and compression-to-ventilation ratio. Each reviewer independently assigned scores based on predefined evaluation criteria. The agreement between the evaluators' ratings was assessed using the intraclass correlation coefficient (ICC).

2.3 Scoring Systems for Assessing Quality and Reliability

The videos were analyzed using three widely utilized and validated scoring systems:

Modified DISCERN Scale (MDS): The MDS was used to assess the reliability of health-related video content. This instrument is derived from the original DISCERN tool, which was developed to evaluate written health information materials, including websites and patient education resources. The modified version consists of five items assessing clarity of aims, reliability of information, objectivity, provision of additional sources, and acknowledgment of areas of uncertainty. Each item is rated on a five-point Likert scale (1–5), yielding a total score ranging from 5 to 25, with higher scores indicating greater reliability. This approach is consistent with previously published YouTube-based health video studies [12,13].

Journal of the American Medical Association (JAMA) scoring system: This quality assessment tool is used to evaluate the reliability and adherence to ethical standards of health-related websites. It specifically measures the accuracy and transparency of online health information. The scoring system is based on four main criteria: authorship, attribution (source referencing), disclosure of conflicts of interest, and currency of information. Each criterion is scored dichotomously (present = 1, absent = 0), resulting in a total score ranging from 0 to 4. A score of 4 indicates

Table 1. CPR-VEI scoring system.

Criterion	Description
1. Preparation	Were the equipment and environment appropriately prepared before the procedure?
2. Positioning	Was the patient positioned correctly for optimal chest compressions and airway management?
3. Chest Compression Technique	Were the compression depth, rate, and continuity appropriate throughout the procedure?
4. Airway Management	Was an appropriate airway technique demonstrated effectively?
5. Defibrillation	Were the shock indications, safety precautions, and defibrillation procedure performed correctly?
6. Team Communication	Were team leadership, command clarity, and coordination adequately demonstrated?
7. Overall Technical Competence	Was the overall CPR algorithm performed fluently and without errors?

Abbreviations: CPR, cardiopulmonary resuscitation; CPR-VEI, cardiopulmonary resuscitation-video evaluation index. Scoring: Each criterion was evaluated using a 6-point ordinal scale (0–5), where 0 = not demonstrated, 1 = very poor, 2 = poor, 3 = acceptable, 4 = good, and 5 = excellent performance. The total CPR-VEI score was calculated by summing the scores of all seven criteria (range: 0–35).

Table 2. CPR-VEI scoring system and rating scale.

Score range	Competency level	Descriptor
1–7	Very Poor	CPR performance does not comply with established standards; essential steps (e.g., compression depth, rate, duration) are performed incorrectly, resulting in no clinical effectiveness.
8–14	Poor	Some CPR components are performed correctly; however, notable deficiencies exist in timing, coordination, or technical accuracy.
15–21	Moderate Competence	Core CPR steps are adequately demonstrated, though deficiencies are observed in aspects such as duration, rhythm recognition, or ventilation technique.
22–28	Good Competence	The procedure is generally accurate and consistent with standard guidelines; minor technical errors are present but do not affect overall clinical adequacy.
29–35	Excellent Competence	CPR performance is complete, timely, well-coordinated, and fully adherent to current resuscitation guidelines, demonstrating outstanding technical proficiency.

excellent quality with full compliance across all criteria, whereas a score of 0 reflects unreliable and non-transparent content [10].

Global Quality Scale (GQS): The GQS is an observer-based, subjective instrument used to evaluate the overall quality and usefulness of health-related videos, websites, and educational materials. It is particularly applied to YouTube™ health videos, patient education resources, and online medical information. The scale employs a five-point Likert-type rating, in which each item is scored from 1 to 5 (1 = poor quality, 5 = excellent quality). When used in conjunction with other scoring systems such as MDS and JAMA, the GQS provides a more comprehensive assessment of the overall quality of online health information [14].

Cardiopulmonary resuscitation-video evaluation index (CPR-VEI): In addition to commonly used validated indices, the CPR-VEI scoring system, a newly developed and study-specific evaluation tool based on key CPR steps defined in current American Heart Association (AHA)/ERC guidelines and structured according to OSATS principles, comprises seven criteria. Each criterion is rated on a scale from 0 to 5, with a maximum total score of 35, and the scoring ranges are classified as follows: 1–7 = very poor; 8–14 = poor; 15–21 = moderate; 22–28 = good; and 29–35 = excellent proficiency. The detailed criteria and scoring rubric for the CPR-VEI system are presented in Tables 1,2.

2.4 Statistical Analysis

The statistical analysis of the data was performed using IBM SPSS Statistics for Windows, version 27.0 (IBM Corp., Armonk, NY, USA). The normality of quantitative variables was assessed using the Kolmogorov–Smirnov test, and all variables were found to exhibit a non-parametric distribution ($p < 0.05$). The Mann–Whitney U test was used for comparisons between two groups, while the Kruskal–Wallis H test was applied for comparisons involving more than two groups. When a statistically significant difference was identified in the Kruskal–Wallis H test, pairwise post hoc comparisons were performed using the Mann–Whitney U test with Bonferroni correction to control for multiple comparisons. Effect sizes for Mann–Whitney U test comparisons were calculated using the formula $r = Z/\sqrt{N}$, where Z represents the standardized test statistic and N represents the total number of observations. Descriptive statistical methods included frequency (n), percentage (%), median, and interquartile range (IQR; 25%–75%). Relationships between variables were evaluated using Spearman's rho correlation analysis, and the coefficients were interpreted as measures of effect size. Inter-rater reliability was assessed using the ICC based on a two-way random-effects model with absolute agreement. As the CPR-VEI scores were derived from multiple independent raters evaluating the same set of videos, inter-rater reliability assessed using the intraclass correlation coefficient (ICC) was considered the most appropriate measure, and Cronbach's al-

pha was therefore not calculated. A p-value < 0.05 was considered statistically significant. Statistical test results (Mann–Whitney U, Z values, and Kruskal–Wallis H statistics) and corresponding effect sizes are presented in the tables.

3. Results

3.1 General Characteristics of the Videos

A total of 120 CPR training videos were analyzed. Of these, 86 (71.7%) were short in duration (1–5 minutes), 28 (23.3%) were of medium duration (5–10 minutes), and 6 (5%) were long (>10 minutes). In terms of upload source, 76 videos (63.3%) were from institutional sources, whereas 44 (36.7%) were uploaded by individual users. Regarding geographic origin, 68 videos (56.7%) were from developed countries (the USA, the UK, Australia, Canada, Belgium, Singapore, and Spain), and 52 (43.3%) from developing countries (Türkiye, India, the Philippines, Pakistan, and Malaysia).

The median and interquartile range (IQR, 25%–75%) values of the quantitative variables are presented in Table 3.

Table 3. Median and interquartile range (IQR 25%–75%) values of quantitative variables.

Variables	n	Median	IQR	
			25%	75%
Number of subscribers	120	50,950	5967	238,000
Number of views	120	90,000	7875	364,000
Number of likes	120	948	53.5	3550
JAMA score	120	2	1	2
CPR-VEI score	120	25	15	30
GQS	120	3	3	4
MDS	120	11	8	13

Wide interquartile ranges reflect the skewed distribution of YouTube™ engagement metrics.

Abbreviations: JAMA, Journal of the American Medical Association; GQS, Global Quality Scale; MDS, Modified DISCERN Scale; IQR, interquartile range.

Based on the predefined CPR-VEI classification, 59 videos (49.2%) were categorized as very poor/poor, 51 (42.5%) as moderate, and only 10 (8.3%) as good/excellent in terms of adherence to guideline-based CPR steps.

3.2 Inter-Rater Reliability

Inter-rater agreement among the evaluators who scored the videos was assessed. The intraclass correlation coefficient (ICC, single measures) was 0.646 [95% confidence interval (CI): 0.415–0.782] for JAMA, 0.962 (95% CI: 0.750–0.986) for CPR-VEI, 0.670 (95% CI: 0.259–0.841) for GQS, and 0.803 (95% CI: 0.433–0.913) for MDS. All ICC values were statistically significant (p < 0.001). Detailed results are presented in Table 4.

3.3 Comparative Analysis by Uploader Type and Country of Origin

Subscriber count, view count, like count, and the JAMA, CPR-VEI, GQS, and MDS were compared according to the country of video upload (developed countries and developing countries). No statistically significant differences were found for any of the variables (p > 0.05 for all). The relationship between quantitative variables and the country of origin of the videos (developed vs developing countries) is presented in Table 5.

View count, JAMA, CPR-VEI, and MDS showed statistically significant differences between institutional and individual videos (p = 0.008, p < 0.001, p = 0.020 and p = 0.031, respectively). Institutional videos had higher view counts as well as higher JAMA, CPR-VEI, and MDS. However, subscriber count, number of likes, and GQS did not differ significantly between institutional and individual uploads (p > 0.05). The relationships between quantitative variables and the type of uploader (institutional vs individual) are presented in Table 6.

3.4 Analysis by Video Duration

View count, number of likes, CPR-VEI, GQS, and MDS showed statistically significant differences across video duration groups (p = 0.014, 0.008, 0.001, 0.002, and p < 0.001, respectively). For the number of views, Bonferroni-adjusted post-hoc analysis demonstrated that intermediate-duration videos had significantly higher view counts than short-duration videos [intermediate duration (ID) > short duration (SD); adjusted p < 0.0167], whereas no significant differences were observed between the remaining pairwise comparisons. Post hoc analysis with Bonferroni correction demonstrated that intermediate-duration videos had significantly higher CPR-VEI, GQS, and MDS scores compared with short-duration videos (p < 0.0167). In addition, long-duration videos showed higher numbers of likes than short-duration videos (p < 0.0167). No significant association was found between video duration and either subscriber count or JAMA score (p > 0.05). The relationship between video duration, number of views, likes, and scoring systems is presented in Table 7.

3.5 Correlation Analysis

A correlation analysis was performed between the CPR-VEI score and the number of likes, as well as the JAMA, GQS, and MDS. A strong positive correlation was identified between CPR-VEI and both GQS and MDS. A moderate positive correlation was found between CPR-VEI and the JAMA score, while a weak positive but statistically insignificant correlation was observed between CPR-VEI score and the number of likes. The correlation coefficients (r) and p-values are presented in Table 8.

4. Discussion

This study systematically evaluated the educational quality and reliability of CPR training videos available

Table 4. Inter-rater reliability analysis using the ICC.

Scale	n	ICC (single measures)	95% confidence interval	ICC (average measures)	*p*-value
JAMA	5	0.646	0.415–0.782	0.901	<0.001
CPR-VEI	5	0.962	0.750–0.986	0.992	<0.001
GQS	5	0.670	0.259–0.841	0.910	<0.001
MDS	5	0.803	0.433–0.913	0.953	<0.001

n represents the number of raters.

Abbreviations: ICC, intraclass correlation coefficient; JAMA, Journal of the American Medical Association; GQS, Global Quality Scale; MDS, Modified DISCERN Scale.

Table 5. Comparison of quantitative variables according to country of origin (developed vs developing countries).

Variables	Group	n	Median (IQR)	U	Z	*p*-value*	r
Number of subscribers	Developed	68	50,000 (5800–238,000)	1657.0	–0.588	0.557	0.05
	Developing	52	53,750 (6085–288,500)				
Number of views	Developed	68	120,500 (9550–410,500)	1384.5	–2.031	0.062	0.19
	Developing	52	58,500 (5250–134,500)				
Number of likes	Developed	68	1050 (53.5–3800)	1625.0	–0.757	0.449	0.07
	Developing	52	778 (55.5–2300)				
JAMA score	Developed	68	2 (1–2)	1535.5	–1.353	0.176	0.12
	Developing	52	1 (1–2)				
CPR-VEI score	Developed	68	25 (20–30)	1586.0	–0.979	0.327	0.09
	Developing	52	25 (15–27)				
GQS	Developed	68	3 (3–4)	1572.0	–1.099	0.272	0.10
	Developing	52	3 (3–4)				
MDS	Developed	68	11 (8–12)	1734.0	–0.181	0.856	0.02
	Developing	52	10.5 (7.5–14)				

*Mann–Whitney U test, Note: $p < 0.05$ indicates statistical significance.

Table 6. Comparison of quantitative variables by uploader type.

Variables	Group	n	Median (IQR)	U	Z	*p*-value*	r
Number of subscribers	Institutional	76	55,100 (7910–238,500)	1585.5	–0.471	0.638	0.04
	Individual	44	45,800 (4490–234,000)				
Number of views	Institutional	76	103,000 (17,000–381,800)	1182.5	–2.666	0.008	0.24
	Individual	44	22,000 (1200–122,000)				
Number of likes	Institutional	76	1300 (128–3800)	1352.0	–1.743	0.081	0.16
	Individual	44	350 (46–2000)				
JAMA score	Institutional	76	2 (1–2)	1066.0	–3.626	<0.001	0.33
	Individual	44	1 (1–2)				
CPR-VEI score	Institutional	76	25 (20–30)	1252.0	–2.324	0.020	0.21
	Individual	44	20 (10–25)				
GQS	Institutional	76	3 (3–4)	1417.5	–1.467	0.142	0.13
	Individual	44	3 (2–4)				
MDS	Institutional	76	11 (8–13)	1278.5	–2.160	0.031	0.20
	Individual	44	10 (6–11)				

*Mann–Whitney U test. Note: $p < 0.05$ indicates statistical significance.

on YouTube™ using widely adopted validated assessment scales and a newly developed CPR-VEI index. Our findings demonstrate that CPR-related content presented in digital environments is not homogeneous. Institutionally produced and professionally structured videos exhibited superior performance in terms of technical accuracy, adherence to resuscitation guidelines, and overall educational adequacy. Significant correlations were identified between CPR-VEI score and the GQS, MDS, and JAMA scores. These findings may indicate that CPR-VEI could serve as a structured tool for the evaluation of video-based CPR education; however, further validation studies are required. Despite a median CPR-VEI score within the moderate range, nearly half of the videos were classified as very poor or poor, indicating substantial variability in adherence to guideline-based CPR steps.

Table 7. Relationship between video duration and quantitative variables.

Variable	n	SD (n = 86) Median (IQR)	ID (n = 28) Median (IQR)	LD (n = 6) Median (IQR)	*p*-value*	H	Post hoc**
Number of subscribers	120	51,000 (5075–238,000)	45,400 (5610–111,000)	460,000 (92,925–698,750)	0.111	4.402	-
Number of views	120	44,000 (5547–249,500)	123,500 (31,250–566,500)	303,500 (83,500–819,000)	0.014	8.589	ID > SD
Number of likes	120	474 (46–2475)	2050 (494–5900)	6500 (1700–16,250)	0.008	9.653	LD > SD
JAMA score	120	1 (1–2)	2 (1–2)	2 (2–3)	0.067	5.407	-
CPR-VEI score	120	20 (15–25)	25 (25–35)	28 (25–30)	0.001	13.405	ID > SD
GQS	120	3 (3–4)	4 (3–5)	4 (3–5)	0.002	12.317	ID > SD
MDS	120	10 (8–11)	12 (11–17)	14 (10–17)	<0.001	16.671	ID > SD

*Kruskal–Wallis H test.

**Pairwise comparisons were performed using the Mann–Whitney U test with Bonferroni correction, and $p < 0.0167$ was considered statistically significant.

Abbreviations: SD, short duration; ID, intermediate duration; LD, long duration.

Table 8. Correlation between CPR-VEI score and video engagement metrics and quality scores.

Parameters	n	Spearman's rho (r)	*p*-value
Number of likes	120	0.156	0.089
JAMA score	120	0.495*	0.001
GQS	120	0.773*	0.001
MDS	120	0.861*	0.001

*Correlation is significant at the 0.05 level (two-tailed). CPR-VEI score was used as the reference variable in the correlation analysis. Spearman's rho correlation coefficient was applied. $p < 0.05$ was considered statistically significant.

Compared with established tools such as JAMA, GQS, and MDS, which primarily assess reliability, transparency, and overall educational quality, the CPR-VEI was specifically designed to evaluate the technical and procedural accuracy of CPR performance based on guideline-defined steps. Therefore, it provides a more structured assessment of practical skill demonstration. However, as a newly developed scoring system, CPR-VEI has not yet undergone external validation and should be interpreted as an exploratory tool. In addition, although it offers a more detailed evaluation of technical performance, it may still be influenced by subjective judgment. It also does not directly measure learning outcomes or clinical effectiveness. In this context, a more detailed examination of the influence of content source and video characteristics on the educational quality of online CPR videos is warranted. Consistent with our subgroup analysis, intermediate-duration videos demonstrated higher CPR-VEI, GQS, and MDS scores compared with short-duration videos, while long-duration videos were associated with higher numbers of likes. The moderate ICC values observed for JAMA and GQS may reflect the more subjective nature of these scales, which rely on qualitative judgments compared with more structured scoring systems.

In our study, no significant differences were observed in subscriber counts, view counts, or like counts, nor in JAMA, CPR-VEI, GQS, and MDS according to the country of video upload. This finding suggests that the geographic location of content creators may not be a primary determinant of the quality and reliability of CPR training videos on YouTube™. Previous studies have shown that online health education materials can reach a global audience and are not limited by regional boundaries. In a systematic review analyzing 22,300 videos, the quality and reliability of health-related content on YouTube™ were found to be more strongly associated with the source of the video, such as academic institutions, healthcare professionals, or individual users, and the uploader's level of expertise than with geographic origin [15]. Similarly, a study comparing coronavirus disease 2019 (COVID-19) related YouTube™ videos uploaded from the USA and South Africa reported regional differences in misinformation levels; however, no direct association was observed between geographic region and overall content quality or reliability [16]. These findings are consistent with our results and suggest that uploader type, source transparency, and content quality may be more influential than country of origin. Moreover, the lack of country-based differences may reflect the global nature of CPR education, in which life-saving content produced in different regions can achieve comparable quality standards. In a study evaluating 121 YouTube™ videos on temporomandibular exercises, it was emphasized that the country of upload, the characteristics of the national healthcare system, or the native language had a smaller-than-expected impact on video quality [17]. In light of our findings and the existing literature, the importance of producing digital health education content within clearly defined universal standards becomes evident. Nevertheless, certain limitations should

be considered when interpreting results based on the country of upload. For example, inaccurate reporting of upload location, the use of English rather than a local language, or multinational content production may affect the validity of country-based classifications. In addition, view and like counts may be influenced by regional dynamics, such as differences between developed and developing countries, and should therefore be interpreted with caution.

The observation that institutionally produced videos demonstrated higher quality and reliability scores than individual content is consistent with findings from previous studies. A large-scale investigation spanning 14 years and encompassing seven databases of YouTube™ medical education videos reported that, although overall content quality was generally low, videos uploaded by academic institutions or healthcare professionals were significantly superior in terms of quality and reliability [6]. Similarly, a 2024 study evaluating 272 videos on trauma management found that content originating from official institutions or healthcare organizations was markedly more reliable and of higher quality than individually uploaded videos [5]. In an analysis of 131 YouTube videos on radical laparoscopic prostatectomy evaluated by three experts, similar findings were reported. In that study, institutionally sourced YouTube videos demonstrated higher JAMA and GQS scores compared to those uploaded by individual users [18]. Another noteworthy finding was the higher view counts observed for institutional videos, suggesting potential advantages in accessibility and visibility and, consequently, a broader educational impact. In contrast, an analysis of 53 YouTube™ videos on chronic pain and peripheral nerve stimulation indicated that popularity does not necessarily correlate with content quality, as viewers often engage with lower-quality material [19]. Viewer preferences, algorithm-driven recommendations, and audiovisual presentation may also influence viewing behavior. Videos containing inaccurate or incomplete information can occasionally achieve high popularity [20]. Nevertheless, the parallel increase in quality scores and view counts for institutional videos in the present analysis suggests that such content may be stronger in both reach and scientific accuracy. Greater technical and academic support, professional production processes, and oversight may contribute to these higher scores. However, uploader type alone does not determine video quality. Factors such as video duration, presentation style, source citation, and content currency also play critical roles [21]. The consistently higher JAMA, CPR-VEI, and MDS associated with institutionally produced videos suggest that these materials may represent a more suitable option for educational purposes.

In our study, view count, number of likes, CPR-VEI, GQS, and MDS scores showed significant differences among videos of varying durations. In particular, intermediate-duration videos had higher CPR-VEI, GQS, and MDS scores compared with short-duration videos, while long-duration videos demonstrated higher numbers of likes than short-duration videos. In contrast, no significant relationship was found between video duration and subscriber count or JAMA score. These findings highlight the impact of video duration on engagement metrics as well as quality and reliability scores of digital health education content. In the literature, video duration has frequently been identified as an important factor influencing the effectiveness of online educational materials. A study conducted in 2021 reported that viewer attention span and learning efficiency are associated with content duration, emphasizing that shorter videos may reduce attention loss [22]. Another study published in 2024 indicated that video length does not have a decisive effect on learning effectiveness, provided that the content is presented in a structured and meaningful manner. Nevertheless, it has also been reported that users tend to prefer shorter videos and that such content may increase viewer engagement [11]. In our study, the observation that intermediate- and long-duration videos achieved higher scores in certain quality and engagement parameters suggests that greater content depth and more comprehensive information delivery may enhance the viewer experience. However, video duration alone does not determine content quality. A study involving neurophysiological measurements in 26 participants demonstrated that video format, narration speed, and content density significantly influence viewer engagement and learning outcomes [23]. When our findings are considered in light of these data, duration alone is not sufficient to ensure high quality. It should be evaluated together with content structure, presentation style, and source reliability. However, it should be emphasized that this association does not indicate a causal effect of video duration on educational quality. It may be confounded by factors such as uploader type, target audience, content scope, and production characteristics. Future studies may benefit from simultaneously examining factors such as video duration, content density, methods of information delivery, and viewer attention dynamics. The observation that intermediate- and long-format videos appear to have advantages in certain quality and engagement metrics suggests that YouTube™ educational content should be optimized. This optimization should be considered in terms of duration in a balanced and purposeful manner.

A weak positive but non-significant association was identified between the CPR-VEI score and the number of likes. Although limited, this finding is noteworthy in the context of digital medical education, as it suggests that popularity indicators, including view and like counts, may not reliably reflect content quality. A systematic review in the field of health informatics emphasized that view and like metrics in educational health videos only marginally represent content quality and are not dependable indicators of educational value [6]. Similarly, another study reported that surface features provided by YouTube™, including view and like counts, are poor proxies for instructional quality, noting that viewer engagement does not necessarily align with educational merit or scientific accuracy [24].

Their study further noted that audience engagement behaviors, such as liking or viewing, may not align with educational merit or scientific accuracy. Rather, these behaviors are often influenced by algorithmic recommendations, presentation style, and audiovisual appeal. In addition, platform algorithms tend to prioritize content based on engagement metrics rather than scientific accuracy. This may increase the visibility of lower-quality videos. Viewer preferences for visually appealing, concise, or entertaining content may further contribute to this mismatch, as such characteristics do not necessarily align with educational depth or guideline adherence. In the present analysis, strong positive correlations were observed between CPR-VEI and both GQS and MDS, while a moderate positive correlation was found with the JAMA score. These findings imply that validated scoring systems assessing structure, accuracy, and reliability provide a more meaningful and objective evaluation of educational quality than popularity metrics. Collectively, the results highlight that engagement indicators like view or like counts should not be used as standalone proxies for educational value. Instead, criteria such as factual accuracy, citation of reliable sources, and overall presentation quality should be prioritized when assessing the educational effectiveness of CPR-related YouTube™ videos.

The 2025 European Resuscitation Council (ERC) Guidelines emphasize the growing importance of digital and technology-supported learning approaches in resuscitation education. The guidelines highlight that online training platforms provide temporal and spatial flexibility, allow individuals to access information at their own pace, and support self-directed learning processes. Moreover, social media-based content is recognized as a potential tool for increasing CPR awareness and promoting public engagement in resuscitation education.

In this context, the ERC recommends that educational institutions develop engaging and interactive digital materials aligned with guideline standards. It also underscores the need for further research investigating the effectiveness of social media-based training and its potential impact on clinical outcomes [25]. In light of these recommendations, our findings suggest that YouTube™-based CPR content may serve as a complementary educational resource, particularly when developed in accordance with evidence-based guidelines and institutional standards. However, given that this study evaluated video characteristics and quality indicators rather than learner outcomes, no direct conclusions can be drawn regarding educational effectiveness. This also applies to skill acquisition and knowledge retention.

This study has several limitations. Due to the dynamic nature of content on the YouTube™ platform, video accessibility and popularity metrics may change over time. Additionally, although the CPR-VEI scoring system was developed to provide a structured and objective evaluation, it has not been externally validated and should be considered an exploratory tool rather than a standardized assessment instrument. Furthermore, variability in video scope (e.g., basic life support versus advanced resuscitation content) may influence scoring and limit comparability across videos. In addition, the absence of multivariable analyses limits the ability to account for potential confounding factors such as uploader type, target audience, content scope, and production characteristics, particularly when interpreting associations related to video duration.

5. Conclusion

Our findings demonstrate that the quality and reliability of CPR training videos on YouTube™ are closely associated with uploader type and content characteristics. Videos produced by institutional sources achieved significantly higher JAMA, CPR-VEI, and MDS scores than those uploaded by individuals, supporting the contribution of professional production and oversight to educational quality. Intermediate-duration videos demonstrated higher educational quality and reliability scores, whereas long-duration videos were associated with higher engagement, particularly in terms of likes; however, popularity indicators such as view and like counts did not reliably reflect educational quality. Ensuring that digital health education materials are developed in accordance with universal standards is important for maintaining global accessibility and scientific consistency. These findings suggest that video-based CPR content may represent a valuable supplementary resource in resuscitation education, particularly when aligned with evidence-based guidelines. However, as this study did not assess learning outcomes, skill acquisition, or clinical performance, no direct conclusions can be drawn regarding educational effectiveness or knowledge retention. Future research should focus on evaluating the impact of video-based CPR training on learning outcomes, skill performance, and clinical practice. Additionally, prioritizing institution-based, evidence-driven educational videos may enhance the quality of available content and support more effective learning environments.

Key Points

• This study demonstrates that the educational quality and reliability of CPR training videos on YouTube™ vary substantially, with nearly half of the videos (49.2%) classified as very poor or poor and only 8.3% demonstrating good or excellent adherence to current resuscitation guidelines.

• Institutionally produced videos consistently showed higher reliability and technical accuracy scores compared with individually uploaded content, suggesting the importance of professional oversight in digital medical education.

• Intermediate-duration videos demonstrated higher educational quality and reliability scores, whereas long-duration videos were associated with higher engagement, particularly in terms of likes.

• Popularity indicators such as view counts and likes did not reliably reflect the educational value or scientific accuracy of CPR videos, highlighting the limitations of using these metrics as quality indicators.

• These findings emphasize the need for healthcare professionals to guide learners toward evidence-based and institutionally developed CPR resources to improve the effectiveness of resuscitation training.

Availability of Data and Materials

The data supporting the findings of this study consist of publicly available YouTube videos. The data used to support the findings of this study are available from the corresponding author upon request.

Author Contributions

Conceptualization: SN and TD; Data curation: SN and BD; Formal analysis: SN and BK; Investigation: SN and AÖ; Methodology: SN and TD; Software: SN and BD; Supervision: SN and AÖ; Validation: SN and BK; Visualization: SN, TD and BD; Writing—original draft: SN, BK and AÖ. All authors contributed to revising the manuscript critically for important intellectual content. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Declaration of Generative AI and AI-Assisted Technologies in Manuscript Preparation

The authors declare that they benefited from artificial intelligence-based tools solely for the purpose of improving grammar and sentence structure. These tools did not contribute to data analysis, interpretation of results, or the generation of scientific content.

References

[1] Perkins GD, Graesner JT, Semeraro F, Olasveengen T, Soar J, Lott C, et al. European Resuscitation Council Guidelines 2021: Executive summary. Resuscitation. 2021; 161: 1–60. https://doi.org/10.1016/j.resuscitation.2021.02.003

[2] Tsao CW, Aday AW, Almarzooq ZI, Anderson CAM, Arora P, Avery CL, et al. Heart Disease and Stroke Statistics-2023 Update: A Report From the American Heart Association. Circulation. 2023; 147: e93–e621. https://doi.org/10.1161/CIR.0000000000001123

[3] Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: A systematic review. Health Informatics Journal. 2015; 21: 173–194. https://doi.org/10.1177/1460458213512220

[4] Topcu H. Evaluation of videos with COVID-19 and CPR content on YouTube. Bagcilar Medical Bulletin. 2023; 8: 57–61. https://doi.org/10.4274/BMB.galenos.2023.2023-01-08

[5] Yazla M, Akyon SH, Aybayar EA, Gedikaslan S, Szarpak L, Turan OF, et al. YouTube as a Source of Information in Trauma Management for ATLS (10th Edition) Guidelines: Evaluation of Trauma Management Videos on YouTube. Emergency Medicine International. 2024; 2024: 7077469. https://doi.org/10.1155/2024/7077469

[6] Helming AG, Adler DS, Keltner C, Igelman AD, Woodworth GE. The Content Quality of YouTube Videos for Professional Medical Education: A Systematic Review. Academic Medicine. 2021; 96: 1484–1493. https://doi.org/10.1097/ACM.0000000000004121

[7] Javidan A, Nelms MW, Li A, Lee Y, Zhou F, Kayssi A, et al. Evaluating YouTube as a Source of Education for Patients Undergoing Surgery: A Systematic Review. Annals of Surgery. 2023; 278: e712–e718. https://doi.org/10.1097/SLA.0000000000005892

[8] Salehpoor-Emran M, Pashaeypoor S, Majdabadi ZA, Böttiger BW, Poortaghi S, Haghani S. The effect of online CPR training on the knowledge and practice of the Red Crescent Student Association Volunteers during the COVID-19 pandemic: A randomized clinical trial study. Resuscitation Plus. 2025; 25: 101010. https://doi.org/10.1016/j.resplu.2025.101010

[9] Aktas BK, Demirel D, Celikkaleli F, Bulut S, Ozgur EG, Kizilkan Y, et al. YouTube™ as a source of information on prostatitis: a quality and reliability analysis. International Journal of Impotence Research. 2024; 36: 242–247. https://doi.org/10.1038/s41443-023-00666-9

[10] Sahin E, Seyyar M. Assessing the scientific quality and reliability of YouTube videos about chemotherapy. Medicine. 2023; 102: e35916. https://doi.org/10.1097/MD.0000000000035916

[11] Seidel N. Short, long, and segmented learning videos: From YouTube practice to enhanced video players. Technology, Knowledge and Learning. 2024; 29: 1965–1991. https://doi.org/10.1007/s10758-024-09745-2

[12] Duran MB, Kizilkan Y. Quality analysis of testicular cancer videos on YouTube. Andrologia. 2021; 53: e14118. https://doi.org/10.1111/and.14118

[13] Kulu B, Uyar SBŞ, Kulu B, Hanci V. Evaluating the quality and reliability of YouTube videos about phenylketonuria. Medicine. 2025; 104: e44309. https://doi.org/10.1097/MD.0000000000044309

[14] Oremule B, Patel A, Orekoya O, Advani R, Bondin D. Quality and Reliability of YouTube Videos as a Source of Patient Information on Rhinoplasty. JAMA Otolaryngology– Head & Neck Surgery. 2019; 145: 282–283. https://doi.org/10.1001/jamaoto.2018.3723

[15] Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. BMC Medical Education. 2022; 22: 382. https://doi.org/10.1186/s12909-022-03446-z

[16] Jung H, Juneja P, Mitra T. Algorithmic behaviors across regions: A geolocation audit of YouTube search for COVID-19 misinformation between the United States and South Africa. Proceedings of the International AAAI Conference on Web and Social Media. 2025; 19: 935–964. https://doi.org/10.1609/icwsm.v19i1.35854

[17] Arikan H, Erol E. Quality and reliability evaluation of YouTube® exercises content for temporomandibular disorders. BMC Oral Health. 2025; 25: 301. https://doi.org/10.1186/s12903-025-05669-0

[18] Akdemir E, Çiçek M, Çakmak BS, Akbulut ML, Buğday MS. Educational Quality of YouTube™ Videos on Laparoscopic Radical Prostatectomy. Journal of Laparoendoscopic & Advanced Surgical Techniques. Part a. 2025; 35: 373–378. https://doi.org/10.1089/lap.2025.0002

[19] Gupta A, Beletsky A, Shen AY, Chin W, Liu C, Reddy R. YouTube as a Source of Medical Information About Peripheral Nerve Stimulation. Neuromodulation. 2025; 28: 1327–1331. https://doi.org/10.1016/j.neurom.2024.09.472

[20] McNamee LA, Shakartzi H, Wasser TE, Li Y, Kim CS. Assessing the accuracy and quality of YouTube videos on early pregnancy loss. Heliyon. 2022; 8: e11947. https://doi.org/10.1016/j.heliyon.2022.e11947

[21] Segado-Fernández S, Herrera-Peco I, Jiménez-Gómez B, Ruiz Núñez C, Jiménez-Hidalgo PJ, Benítez de Gracia E, et al. Realfood and Cancer: Analysis of the Reliability and Quality of YouTube Content. International Journal of Environmental Research and Public Health. 2023; 20: 5046. https://doi.org/10.3390/ijerph20065046

[22] Krumm IR, Miles MC, Clay A, Carlos Ii WG, Adamson R. Making Effective Educational Videos for Clinical Teaching. Chest. 2022; 161: 764–772. https://doi.org/10.1016/j.chest.2021.09.015

[23] Lackmann S, Léger PM, Charland P, Aubé C, Talbot J. The Influence of Video Format on Engagement and Performance in Online Learning. Brain Sciences. 2021; 11: 128. https://doi.org/10.3390/brainsci11020128

[24] Bitzenbauer P, Hofler S, Veith JM, Winkler B, Zenger T, Kulgemeyer C. Exploring the relationship between surface features and explaining quality of YouTube explanatory videos. International Journal of Science and Mathematics Education. 2024; 22: 25–48. https://doi.org/10.1007/s10763-022-10351-w

[25] Greif R, Lauridsen KG, Djärv T, Ek JE, Monnelly V, Monsieurs KG, et al. European Resuscitation Council Guidelines 2025 Executive Summary. Resuscitation. 2025; 215: 110770. https://doi.org/10.1016/j.resuscitation.2025.110770