

Measurement of the Reliability and Quality of Online Surgery Videos with Artificial Neural Networks

İdris Kurtuluş

University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Clinic of General Surgery, İstanbul, Turkey

ABSTRACT

Introduction: The effect of the internet on education has started to increase more and more with the Coronavirus disease-2019 (COVID-19) pandemic. Although many researchers have used online resources during the pandemic, their quality and reliability and the way these two aspects could be evaluated online has remained a problem. This study aims to measure the reliability and quality of online videos about inguinal hernias by using the DISCERN questionnaire and to see if the quality and reliability of these online videos are determined accurately and fast with artificial neural networks (ANNs) by teaching them some easily accessed variables about the videos.

Methods: A total of 30 online videos searched on Google with the keywords of “TEP,” and “totally ekstraperitoneal inguinal hernia repair” from February 15 to March 1, 2021 with the approval of the University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital Ethical Committee were included in this research (approval number: 303, date: 29.12.2021). The videos were found using the “videos” tab of Google. The DISCERN questionnaire was applied to the videos and the results of the questionnaire were tried to be estimated with ANNs by teaching them some easily accessed variables about the videos. The results of the questionnaire and results of the predictions were compared.

Results: A total of 30 videos were evaluated. Benefitting from the scores of the DISCERN questionnaire, a total of three groups were formed with K-means clustering analysis. The scores of the low-quality videos were between 0 and 26.50, of the medium quality videos between 26.50 and 34.9, and of the high-quality ones between 34.9 and 48. In determining the video quality in the ANNs estimation model, “the number of likes” (a video received) was found to be the variable with the highest effect (on the model) with an importance coefficient of 0.245. Its normalized importance was found to be 100%. “Country” with an importance coefficient of 0.060 was found to be the variable with the lowest effect (on the model). Its normalized importance was found to be 24.4%. The median value of the scores obtained from the DISCERN questionnaire was 28, while the median value of the DISCERN scores estimated with the ANNs was 27.50. No statistically significant difference was observed between the distribution of the estimated scores and the scores obtained from the DISCERN questionnaire ($p=0.314$). Additionally, there was no statistically significant difference between the video groups formed according to the scores of the DISCERN questionnaire, and the video groups formed according to the DISCERN scores that were estimated with the ANNs ($p=0.771$). A total of 20 videos were found to be low quality with the DISCERN questionnaire, while 19 videos were estimated to have low quality with the ANNs. The number of medium-quality videos was six according to the DISCERN questionnaire and five according to the ANNs. There were a total of four high-quality videos in the DISCERN questionnaire and six in the ANNs. It was observed that the quality group of 86.6% of the videos (26 videos) was predicted accurately with the ANNs.

Conclusion: The quality and reliability of most TEP videos on the Internet were quite low. The instruments in the literature make searches retrospectively, and using them is quite time consuming. ANNs seem to be quite successful in estimating the reliability and quality of online videos. The reliability and the quality of the videos could be shown to users fast with online electronic labels developed thanks to the ANNs.

Keywords: DISCERN, totally ekstraperitoneal hernia repair, quality and the reliability of videos, TEP, artificial neural networks

Introduction

It has been more than a year since the start of the Coronavirus disease-2019 (COVID-19) pandemic. The beginning of the vaccination process offers hopes about the issue yet concerns about it continue to increase as the virus is mutating. This recent development has adversely affected many industries/sectors (1). Serious problems have started to

emerge in the area of education. The negative effect of the COVID-19 pandemic on medical education has started to be experienced to varying degrees. Complications have also surfaced in the education of general surgery. Although academics have tried to narrow the gap the pandemic has created in education by setting up online courses and congresses, debates about the efficiency of online education are continuing since



Address for Correspondence: İdris Kurtuluş MD, University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Clinic of General Surgery, İstanbul, Turkey
Phone: +90 532 170 70 47 **E-mail:** idriskurtulus@gmail.com **ORCID ID:** orcid.org/0000-0001-9101-6363

Cite this article as: Kurtuluş İ. Measurement of the Reliability and Quality of Online Surgery Videos with Artificial Neural Networks. İstanbul Med J 2022; 23(2): 79-84.

Received: 09.12.2021
Accepted: 09.03.2022

it is not known how to standardize it. However, it seems it is high time a reform is done in surgical education. Using teletechnologies could make a revolution in surgical education as much as using augmented reality, which enables students to observe surgeons while they perform operations and to have interactions from distant locations for more extensive experience, could do (2).

In their article, which focuses on medical education in Turkey during the COVID-19 pandemic, Tokuç and Varol (3) have stated that medical education is an area open to changes and developments as other fields are, and educators should analyze the effects of current changes (on education) with their students to determine new education principles and applications. Websites that offer various contents and information have already become a part of our lives with the prevailing of the internet quick worldwide. Accessing online materials has become particularly significant for surgeons due to such technological developments. YouTube, TVASURG and WebSurg are the best known websites (among health professionals), and some of them are open access websites that allow their users to share and watch videos at academic levels (4). However, no instruments measure the standardization of the parameters of the online websites, such as the reliability, quality and certification of their contents at an instant and fast. Researchers have only been able to conduct retrospective analyzes of the issue with different methods.

Our purpose in this study is to measure the reliability and quality of online videos about inguinal hernias with the DISCERN questionnaire and to see whether the quality and reliability of the online videos are determined accurately and fast with artificial neural networks (ANNs) by teaching them some easily accessed variables about the videos.

Methods

A total of 30 online videos searched on Google with the keywords of “TEP,” and “totally ekstraperitoneal inguinal hernia repair” from February 15 to March 1, 2021 with the approval of the University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital Ethical Committee were included in this research (approval number: 303, date: 29.12.2021). The videos were found using the “videos” tab of Google. The videos included in the study are videos publicly accessible by web browsers. A DISCERN questionnaire was applied to the videos and the results of the questionnaire were tried to be estimated with ANNs by teaching them some easily accessed variables about the videos. The results of the questionnaire and results of the predictions were compared. Online videos with the purpose of advertisement were excluded from the study.

The countries where the videos were uploaded, the people who uploaded the videos, the period during which the videos were available, their broadcast language, the duration of the videos, and the number of times they were watched and liked were noted down.

The DISCERN questionnaire was applied to all the videos by the same general surgeon. The HONcode certification of the videos was taken into consideration and their Video Power Index (VPI) was calculated.

The DISCERN scores were calculated for each video and a total of three groups were formed with the K-means cluster analysis by using them. It was determined which videos belonged to which group. Then, an estimation model of ANNs was established. In this model, the scores of the DISCERN

questionnaire were tried to be estimated according to the countries where the videos were uploaded, the people who uploaded the videos, the period the videos was available online, their broadcast language, their duration and the number of times the videos were watched and liked, as well as their HONcode certification and VPI. Again, the DISCERN scores obtained with the estimation model of ANNs were grouped with K-means clustering analysis. The clusters obtained according to the results of the DISCERN questionnaire and the clusters predicted according to the estimated DISCERN questionnaire scores were compared. It was not observed whether there was a difference between them.

In the ANN model, 73.3% of the videos were planned to be used for training, 20% for testing and 6.7% as holdout (In our study, we used the Bernoulli distribution to separate our dataset. According to this distribution, 24 videos were selected for training and 6 videos for testing. To use hold-out in our study, we determined two of the 24 training videos as hold-out). Learning type was specified as online. The HONcode certification, the source of the videos / (the people who uploaded them), their language and countries were used as factors in the input layer, and the year of the videos, their VPI, and the number of times they were liked and disliked and watched and were used as scalar variables. The scores of the DISCERN questionnaire were used as dependent variables. (As the maximum DISCERN score is 80), the highest score of the random number generator was decided to be 80 to prevent it from forming an initial value from a random point in SPSS and to standardize the results before starting to form the estimation model of the ANNs. The automatic number appointment function of the random number generator was turned off.

The DISCERN Questionnaire

It was developed by researchers from the University of Oxford to determine the quality of health information and suggested options of treatment (5,6). The scores of the questionnaire ranged between 0 and 80. It is made up of three parts: the first part evaluates the reliability of a publication (questions; 1-8), the second part investigates the quality of the information about the treatment options (questions; 9-15), and the third part has a single question (question; 16) and evaluates the publication in its general aspects.

The first section uses a series of widely accepted evaluation criteria with respect to consistency, clarity, relevancy, objectivity, certainty and the references provided. The second section is quite specific to the DISCERN instrument and evaluates the quality of information about treatment options.

Web Certificate (HON=Health on NET)

Many websites offer certificates developed for the reliability of health information. Certifications are aimed at standardizing the reliability of information. The most widely used one is Health on the Net Foundation's HONcode certification. Health on the Net Foundation is an internationally accepted non-profit organization founded in 1995. Currently, more than 8000 websites in 102 countries use the HONcode certification. It provides its users with an electronic label called “HON label”, and they load it on their browsers. The label easily shows if a website has the HON certification when it is clicked on (7).

Video Power Index

This is a scale that measures the popularity of a video. It has two basic parameters. They are about a video's being watched and liked. It was first used by Erdem and Karaca (8) It is calculated with the formula of "VPI=like ratio X view ratio/100."

Statistical Analysis

Statistical analysis was done with IBM SPSS V 25. Kolmogorov-Smirnov and Shapiro-Wilk tests were used to verify the normalization of distribution. Mann-Whitney U test or Student's t-test was used for the comparisons between the groups. Chi-square test was used to compare the categorical information. Fisher's exact test and Pearson's chi-square test were used to evaluate categorical information. Questionnaire scores were grouped with K-means cluster analysis. An estimation model was established with the model of ANNs. The hold-out method was used for model selection and hyperparameter validation. The variables' normalized importance coefficients were calculated. The results were evaluated with 95% confidence interval and $p < 0.05$ was accepted to be statistically significant.

Results

A total of 30 videos were evaluated. Table 1 shows the results of the descriptive statistics of the videos. A total of three groups were formed with the K-means clustering analysis according to the scores obtained from the DISCERN questionnaire. The videos with low quality were rated between 0 and 26.50, the videos with medium quality between 26.50 and 34.9, and the ones with high quality between 34.9 and 48. Of the videos, 66.6% were seen to be low quality, 20% medium quality and 13.4% low quality. Details of the data are shown in Table 2.

New DISCERN scores determined using ANNs were calculated using the scores obtained from the DISCERN questionnaire. In the ANN model, 73.3% of the videos were planned to be used for training, 20% for testing and 6.7% as holdout. A hidden layer was seen and eight subcategories were observed in it. The synaptic network of the ANNs is shown in Figure 1. Input and output data of the ANN model are shown in Table 3.

In determining the video quality in the ANNs estimation model, "the number of likes" was found to be the variable with the highest effect (on the model) with an importance coefficient of 0.245. Its normalized importance was observed to be 100%. "Country" was found to be the variable with the lowest effect (on the model) with an importance coefficient of 0.060 in the estimation model. Its normalized importance was observed to be 24.4%. Details of the data are shown in Table 4. The median value of the scores obtained from the DISCERN questionnaire was 28, and the median value of the DISCERN scores predicted by the ANNs was 27.50. There was no statistically significant difference between the distributions of the estimated scores and the scores obtained from the DISCERN questionnaire ($p=0.314$). Details of the data are shown in Table 5.

No statistically significant difference was observed between the groups formed according to the scores of the DISCERN questionnaire and the groups formed with K-means cluster analysis according to the DISCERN scores that were estimated with the ANNs ($p=0.771$). Of the DISCERN

Table 1. Descriptive statistics of the videos

| | | n=30 |
|--------------------------------------|-------------------------|----------------------|
| Country (frequency, %) | India | 8 (26.7%) |
| | Turkey | 3 (10.0%) |
| | Australia | 7 (23.3%) |
| | America | 5 (16.7%) |
| | Brazil | 4 (13.3%) |
| | Unknown | 3 (10.0%) |
| Language (frequency, %) | English | 11 (36.7%) |
| | Turkish | 3 (10.0%) |
| | Spanish | 3 (10.0%) |
| | Without sound | 13 (43.3%) |
| Video sources (frequency, %) | Academics | 9 (30.0%) |
| | Medical specialists | 7 (23.3%) |
| | Patients | 4 (13.3%) |
| | Unknown | 10 (33.3%) |
| HONcode certification (frequency, %) | Available/unavailable | 2 (6.7%); 28 (93.3%) |
| Number of watches | Minimum/median/maximum | 270/3975/44357 |
| Number of likes | Minimum/median/maximum | 9/52/451 |
| Number of dislikes | Minimum/median/maximum | 0/8/23 |
| VPI | Minimum/median/maximum | 20718/514602/7145699 |
| DISCERN scores | WMinimum/median/maximum | 21/28/48 |

VPI: Video Power Index

Table 2. Distribution of the videos made with K-means cluster analysis according to the DISCERN scores

| | Low quality (0-26.50) | Medium quality (26.5-34.9) | High quality (34.9-48) |
|----------------------|-----------------------|----------------------------|------------------------|
| Video numbers (n=30) | 20 (66.6%) | 6 (20%) | 4 (13.4%) |

questionnaire, 20 videos were found to have low quality. According to the ANN results, 19 videos were observed to have low quality. The number of videos with medium quality was six according to the results of the DISCERN questionnaire, while it was five according to the results of the ANNs. Four videos had high quality according to the results of the DISCERN questionnaire and this number was six according to the ANNs. We observed that the quality groups of the videos were predicted to be 86.6% accurate (26 videos) with the ANNs. Details of the data are shown in Table 6.

Discussion

In surgical education about hernia surgeries, a comparison between traditional methods of surgery and endoscopic methods reveals that there are theoretical and practical difficulties in learning and applying the latter. For this reason, learning and teaching of innovative methods of surgery tend to continue even after the training of the assistants. General surgeons who are interested in this field have been trying to learn about these techniques from various sources, including online

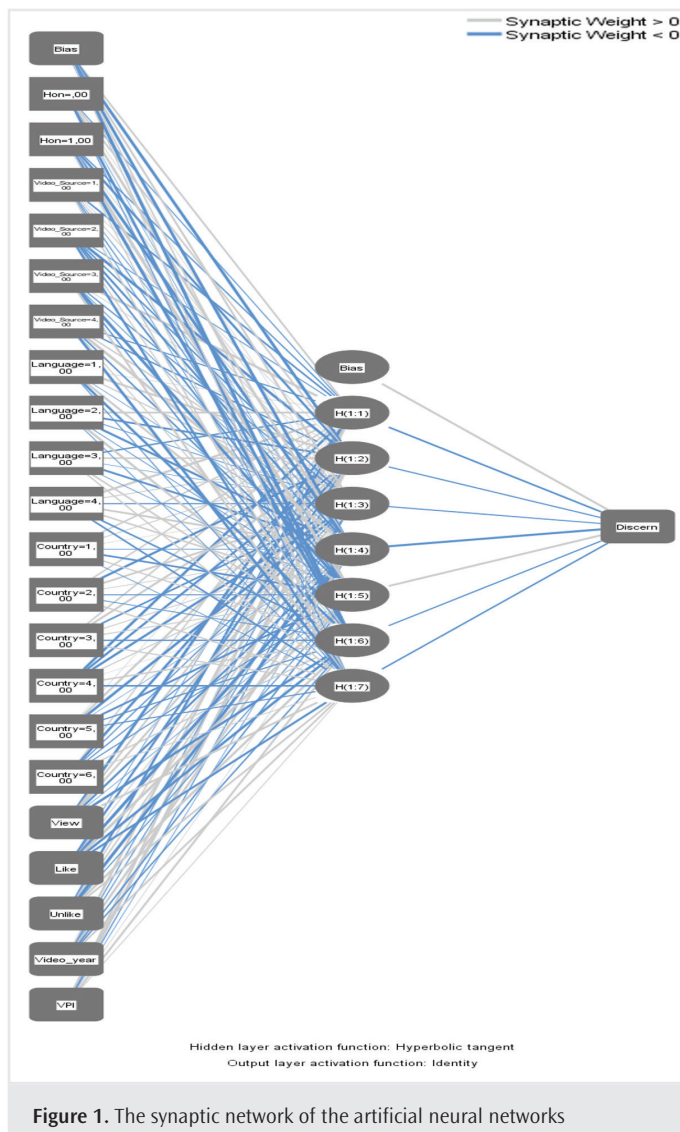


Figure 1. The synaptic network of the artificial neural networks

tools, courses and conferences. However, the COVID-19 pandemic still continues to affect the world although more than a year has passed since it starts. The beginning of the vaccination process offers hopes about the issue yet concerns about it continue to increase as the virus is mutating. This recent development has adversely affected many industries/sectors (1). Online education has become more widespread in many countries due to the pandemic. Such an orientation has frequently brought forth the questions of “What are the quality principles for the reliability of health websites?” and “How should they be measured?”

The DISCERN questionnaire is a website evaluation tool whose reliability, credibility and internal consistency have been regarded as positive in many studies (9,10). Users of the questionnaire are not limited to specialists and scientists, instead, any person with health literacy could use it to evaluate the quality of a website. The instrument is also popular among researchers and it is the most widely used tool in evaluating the quality of information on health (11). There were 16 questions on the questionnaire. Each question had a score from 1 to 5. However, applying the questionnaire is not easy. It was seen in our study that applying the questionnaire takes 30 min for each video, regardless of the duration of

Table 3. Data about artificial neural networks model

| | | | |
|---------------------------------|---------------------------------------|--------------------|-----------------------|
| Sample (n=30) | Training | 22 (73.3%) | |
| | Test | 6 (20.0%) | |
| | Holdout | 2 (6.70%) | |
| | Learning type | Online | |
| Input layer | Factors | 1 | HONcode certification |
| | | 2 | Video source |
| | | 3 | Language |
| | | 4 | Country |
| | Covariates | 1 | Video year |
| | | 2 | VPI |
| | | 3 | Number of watches |
| | | 4 | Number of likes |
| | | 5 | Number of dislikes |
| | Number of units | | 21 |
| Rescaling method for covariates | | Normalized | |
| Hidden layer(s) | Number of hidden layers | 1 | |
| | Number of units in hidden layer | 8 | |
| | Activation function | Hyperbolic tangent | |
| Output layer | Dependent variable | 1 | DISCERN |
| | Number of units | 1 | |
| | Rescaling method for scale dependents | Standardized | |
| | Activation function | Identity | |
| Error function | | Sum of squares | |

VPI: Video Power Index

Table 4. Importance of variables in the estimation of video quality according to the ANNs

| | Importance coefficient | Normalized importance |
|--------------------|------------------------|-----------------------|
| HON certification | 0.124 | 50.5% |
| Video source | 0.101 | 41.3% |
| Language | 0.071 | 29.0% |
| Country | 0.060 | 24.4% |
| Video year | 0.141 | 57.4% |
| VPI | 0.113 | 46.2% |
| Number of watches | 0.068 | 27.8% |
| Number of likes | 0.245 | 100.0% |
| Number of dislikes | 0.077 | 31.6% |

VPI: Video Power Index, ANNs: Artificial neural networks

the video. It is unreasonable for researchers to devote 30 min to each video they watch. It is undebatable that there is a need for an online system that could provide its users with ideas about the reliability and quality of the contents of a website.

Measuring parameters for such a system must include parameters that could be used and interpreted by machines fast as data on the Internet. How many times a video is watched, its language, the country it is loaded on the internet, the source it is loaded from, and its web certification are easily accessed parameters. There are also metrics such

Table 5. Distribution of DISCERN scores and the DISCERN scores estimated according to the ANNs (p=0.314)

| | Minimum | Median | Maximum |
|---|---------|--------|---------|
| DISCERN scores | 21.00 | 28.00 | 48.00 |
| DISCERN scores estimated according to the ANN | 20.90 | 27.50 | 50.38 |

ANNs: Artificial neural networks

Table 6. Estimated distribution of the videos according to the DISCERN and ANNs (p=0.771)

| | Low quality (0-26.50) | Medium quality (26.5-34.9) | High quality (>34.9) |
|---------------------------------------|-----------------------|----------------------------|----------------------|
| Number of videos according to DISCERN | 20 (66.6%) | 6 (20%) | 4 (13.4 %) |
| Number of videos according to ANNs | 19 (73.2%) | 5 (13.4%) | 6 (13.4 %) |

ANNs: Artificial neural networks

as VPI that could evaluate the popularity of online videos (8). The most common tool that could enable machines to learn with these metrics is ANNs.

ANNs are parallel and distributed knowledge processing structures which are developed with inspiration from the human brain, connected to each other with weight connections, and composed of processing units, each of which has its own memory. In other words, they are computer programs that imitate biological neural networks.

The first model for an ANN was developed in 1943 by neurophysiologist Warren Sturgis McCulloch and mathematician Walter Pitts (12). The ANNs are frequently used in preparing estimations, categorizations and early projections. Many researchers from the field of medicine use the ANNs as well (13-15). The basic structure of the neurons of the ANNs is made up of inputs, weights, summation functions, activation functions and outputs.

In evaluating the educational quality of a video, the Global Quality Scale developed by Mutter et al. (16) and the Journal of American Medical Association benchmark criteria proposed by Bernard et al. (5) are also used. In this study, the DISCERN questionnaire was preferred. We first investigated the videos that are about the surgeries performed with the total ekstraperitoneal (TEP) approach with the DISCERN questionnaire, and we later investigated whether the quality and reliability of these videos could also be estimated via ANNs.

In the first part of the study, the DISCERN questionnaire was applied to the videos and videos were divided into groups according to the results. The results of the DISCERN questionnaire showed that 66.6% of the videos had low quality, 20% had medium quality and 13.4% had good quality. As the second step, the DISCERN scores were appointed as dependent variables to the ANNs. The HON certification, the number of times videos were liked and disliked and watched, VPI, the sources of the videos, their languages, countries and years were used as inputs.

Prasanth et al. (17) used the HONcode certification to evaluate the quality of online information about testicular cancer yet they could not

make a connection between the scores of the DISCERN questionnaire and the HON certification. Kartal and Kebudi (6) noticed that there is a positive correlation between VPI and (the scores of) the DISCERN questionnaire, and the number of times videos were watched and (the scores of) the DISCERN questionnaire, and there is a weak negative correlation between the years of the videos and the DISCERN (scores). They stated that there is no statistically significant difference between the DISCERN scores of the videos with respect to their sources and countries. In their studies on TEP videos on Youtube, Kanlioz and Ekici (18) found a positive correlation between the VPI scores and the number of times a video is watched, and the DISCERN scores and the number of times a video is liked. Additionally, they found that the DISCERN scores of the videos uploaded by academics are higher although there is no statistically significant difference between the two factors.

In our study, the variable with the highest importance coefficient was “the number of times a video was liked,” which was a variable used in the ANNs. the year of the videos, if they had the HON certification, their VPI scores, the source of the videos, the number of times they were liked, their language, the number of times they are watched and their countries had, respectively, positive effect in predicting the DISCERN scores accurately.

In the basic structure of the ANNs’ model, a mathematical function is formed using some learning algorithms with the multiple connections between inputs and outputs. Each unit is linked to each other with weights according to its importance in the dataset. These weights show the connection strength of the two units each.

In our study, there was no statistical difference between the DISCERN scores obtained through the ANNs and scores obtained via the DISCERN questionnaire itself (p=0.314). In the groupings of the video quality estimated with the ANNs and with the DISCERN questionnaire, four videos went to different groups.

However, there was no statistical difference between the video groups formed according to the scores of the DISCERN questionnaire and the video groups formed with K-means clustering analysis according to the DISCERN scores estimated with the ANN (p=0.771). In the DISCERN questionnaire 66.6% of the videos were seen to have low quality while this percentage was 73.3 in the ANNs. Twenty percent of the videos had medium quality according to the DISCERN questionnaire while this rate was 13.4 in the ANNs. In the DISCERN questionnaire 13.4% of the videos had high quality, and in the ANNs 13.4% of them had high quality. It was observed that with the ANNs, the quality category of 86.6% of the videos (26 videos) was estimated accurately.

The ANNs are made up of many neurons and they could conduct complex tasks simultaneously. They do deep learning. They could easily solve problems with linear relations or without them. They learn with machine learning and could make very logical decisions when they encounter similar situations. They could even make generalizations about the issues they have not seen before. In our study, it was observed that with the ANNs, the quality category of 86.6% of the videos (26 videos) was estimated accurately. It is sure that further studies need to be carried out in this field. Still, it is anticipated that fruitful results could be obtained by teaching ANNs other instruments of measurement

frequently used in the literature. It is hoped that users will be offered information about the quality and reliability of the videos they will watch with the creation of electronic labels such as the HON label by using the ANNs.

Study Limitations

The most important limitation of this study was that it was retrospective and the number of evaluated videos was low.

Conclusion

The quality and reliability of most online videos about inguinal hernia are quite low. The instruments used in the literature make research retrospectively and are time consuming. However, it seems that ANNs are quite successful in estimating the reliability and quality of the online videos. The reliability and quality of the videos could be easily shown to their audiences with online labels developed with the ANNs.

Ethics Committee Approval: Our study was approved by the Ethics Committee of University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital (approval number: 303, date: 29.12.2021).

Informed Consent: Retrospective study.

Peer-review: Externally and internally peer-reviewed.

Financial Disclosure: The author declared that this study received no financial support.

References

1. Ahmed H, Allaf M, Elghazaly H. COVID-19 and medical education. *Lancet Infect Dis* 2020; 20: 777-8.
2. Greenfield MJ, Luck J, Billingsley ML, Heyes R, Smith OJ, Mosahebi A, et al. Demonstration of the Effectiveness of Augmented Reality Telesurgery in Complex Hand Reconstruction in Gaza. *Plast Reconstr Surg Glob Open* 2018; 6: e1708.
3. Tokuç B, Varol G. Medical education in Turkey in time of COVID-19. *Balkan Med J* 2020; 37: 180-1.
4. Ferhatoglu MF, Kartal A, Ekici U, Gurkan A. Evaluation of the Reliability, Utility, and Quality of the Information in Sleeve Gastrectomy Videos Shared on Open Access Video Sharing Platform YouTube. *Obes Surg* 2019; 29: 1477-84.
5. 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: 2070-7.
6. Kartal A, Kebudi A. Evaluation of the Reliability, Utility, and Quality of Information Used in Total Extraperitoneal Procedure for Inguinal Hernia Repair Videos Shared on WebSurg. *Cureus* 2019; 11: e5566.
7. HON code: (date of access: 24.02.2022): <https://www.hon.ch/HONcode/>
8. Erdem MN, Karaca S. Evaluating the accuracy and quality of the information in Kyphosis videos shared on YouTube. *Spine (Phila Pa 1976)*. 2018; 43: E1334-9.
9. Griffiths KM, Tang TT, Hawking D, Christensen H. Automated assessment of the quality of depression websites. *J Med Internet Res* 2005; 7: e59.
10. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* 2002; 287: 2691-700.
11. Zhang Y, Sun Y, Xie B. Quality of health information for consumers on the web: a systematic review of indicators, criteria, tools, and evaluation results. *J Assoc Inf Sci Technol* 2015; 66: 2071-84.
12. Elmas Ç. Artificial neural networks (Theory, Architecture, Training, Implementation), 1st Edition, Ankara: Seçkin Press; 2003 (In Turkish).
13. Renganathan V. Overview of artificial neural network models in the biomedical domain. *Bratisl Lek Listy* 2019; 120: 536-40.
14. Kapoor R, Walters SP, Al-Aswad LA. The current state of artificial intelligence in ophthalmology. *Surv Ophthalmol* 2019; 64: 233-40.
15. Currie G, Hawk KE, Rohren E, Vial A, Klein R. machine learning and deep learning in medical imaging: intelligent imaging. *J Med Imaging Radiat Sci* 2019; 50: 477-87.
16. Mutter D, Vix M, Dallemagne B, Perretta S, Leroy J, Marescaux J. WebSurg: An innovative educational Web site in minimally invasive surgery--principles and results. *Surg Innov* 2011; 18: 8-14.
17. Prasanth AS, Jayarajah U, Mohanappirian R, Seneviratne SA. Assessment of the quality of patient-oriented information over internet on testicular cancer. *BMC Cancer* 2018; 18: 491.
18. Kanlıoz M, Ekici U. Reliability and educational features of YouTube videos about hernia operations performed using laparoscopic TEP method. *Surg Laparosc Endosc Percutan Tech* 2020; 30: 74-8.