Universitat de València

Facultat de Fisioteràpia

Programa de Doctorado en Fisioterapia

Doctoral thesis by compendium of publications with International Mention

"QUALITY ANALYSIS OF AVAILABLE INFORMATION FOR PATIENTS WITH DIFFERENT PATHOLOGIES ON SOCIAL MEDIA VIDEO PLATFORMS."

Author

Álvaro Manuel Rodríguez Rodríguez

Supervisors

Prof. María Blanco Díaz

Prof. Jose Casaña Granell

Prof. Sergio Hernández Sánchez

Valencia, 2022

Universitat de València

Facultat de Fisioteràpia

Programa de Doctorado en Fisioterapia

Doctoral thesis by compendium of publications with International Mention

"QUALITY ANALYSIS OF AVAILABLE INFORMATION FOR PATIENTS WITH DIFFERENT PATHOLOGIES ON SOCIAL MEDIA VIDEO PLATFORMS."

Author

Álvaro Manuel Rodríguez Rodríguez

Supervisors

Prof. María Blanco Díaz

Prof. Jose Casaña Granell

Prof. Sergio Hernández Sánchez

Valencia, 2022

La Dra. María Blanco Díaz, Profesora Ayudante Doctor de la Universidad de Oviedo, adscrita al Departamento de Cirugía y Especialidades Médico-Quirúrgicas,

El Dr. Jose Casaña Granell, Profesor Titular de la Universidad de Valencia, adscrito al Departamento de Fisioterapia,

El Dr. Sergio Hernández Sánchez, Profesor Contratado Doctor de la Universidad Miguel Hernández, adscrito al Departamento de Fisioterapia,

Hacen costar:

> Que el trabajo presentado como Tesis Doctoral por D. Álvaro Manuel Rodríguez Rodríguez, titulado "**QUALITY ANALYSIS OF AVAILABLE INFORMATION FOR PATIENTS WITH DIFFERENT PATHOLOGIES ON SOCIAL MEDIA VIDEO PLATFORMS**", ha sido realizado bajo nuestra dirección para optar al grado de Doctor por la Universitat de València.
>
> Que habiéndose concluido, y reuniendo a su juicio las condiciones de originalidad y rigor científico necesarias, se autoriza su presentación a fin de que pueda ser defendido ante el tribunal correspondiente.

Y para que así conste, expiden y firman la presente certificación en València, a 05 de Diciembre de 2022.

BLANCO DIAZ MARIA - 09416087W Firmado digitalmente por BLANCO DIAZ MARIA - 09416087W Fecha: 2022.12.05 15:47:53 +01'00'	Nombre de reconocimiento (DN): cn=JOSE CASAÑA GRANELL - NIF:20154083B, serialNumber=20154083B, givenName=JOSE, sn=CASAÑA GRANELL, ou=CIUDADANOS, o=ACCV, c=ES Fecha: 2022.12.05 15:58:49 +01'00'	HERNANDEZ SANCHEZ, SERGIO (FIRMA) Firmado digitalmente por HERNANDEZ SANCHEZ, SERGIO (FIRMA) Fecha: 2022.12.05 17:26:03 +01'00'
Fdo.: Dra. María Blanco Díaz (Directora)	Fdo.: Dr. Jose Casaña Granell (Director y tutor)	Fdo.: Dr. Sergio Hernández Sánchez (Director)

La Dra. María Blanco Díaz, Profesora Ayudante Doctor de la Universidad de Oviedo, adscrita al Departamento de Cirugía y Especialidades Médico-Quirúrgicas,

El Dr. Jose Casaña Granell, Profesor Titular de la Universidad de Valencia, adscrito al Departamento de Fisioterapia,

El Dr. Sergio Hernández Sánchez, Profesor Contratado Doctor de la Universidad Miguel Hernández, adscrito al Departamento de Fisioterapia,

Hacen costar:

Que la tesis doctoral realizada por D. Álvaro Manuel Rodríguez Rodríguez, y presentada a continuación, sigue el modelo de compendio de artículos científicos con Mención Internacional.

Que el autor de la presente tesis es primer firmante de los tres artículos que se presentan. La contribución del doctorando en los artículos científicos se ha hecho patente en las diferentes fases del proyecto, como son el diseño de la misma y la ejecución de ésta, así como la redacción de los artículos científicos.

Y para que así conste, expiden y firman la presente certificación en València, 05 de Diciembre de 2022.

BLANCO DIAZ MARIA - 09416087W
Firmado digitalmente por BLANCO DIAZ MARIA - 09416087W
Fecha: 2022.12.05 15:48:24 +01'00'

Fdo.: Dra. María Blanco Díaz
(Directora)

Nombre de reconocimiento (DN): cn=JOSE CASAÑA GRANELL - NIF:20154083B, serialNumber=20154083B, givenName=JOSE, sn=CASAÑA GRANELL, ou=CIUDADANOS, o=ACCV, c=ES
Fecha: 2022.12.05 15:59:10 +01'00'

Fdo.: Dr. Jose Casaña Granell
(Director y tutor)

HERNANDEZ SANCHEZ, SERGIO (FIRMA)
Firmado digitalmente por HERNANDEZ SANCHEZ, SERGIO (FIRMA)
Fecha: 2022.12.05 17:27:06 +01'00'

Fdo.: Dr. Sergio Hernández Sánchez
(Director)

"La Calavera Garbancera" (La Catrina)

José Gualdalupe Posada, 1913

"Feet, what do I need you for when I have wings to fly?"

Frida Kahlo

IDCCultura
Instituto Internacional de Derecho Cultural y Desarrollo Sustentable

EL INSTITUTO INTERNACIONAL DE DERECHO CULTURAL Y DESARROLLO SUSTENTABLE IIDECUL DESUS, A.C. IDC CULTURA- MÉXICO

EXPIDE EL PRESENTE CERTIFICADO DE ESTANCIA DE INVESTIGACIÓN A

Álvaro Manuel Rodríguez Rodríguez

Por el trabajo desempeñado en el marco de su Tesis Doctoral

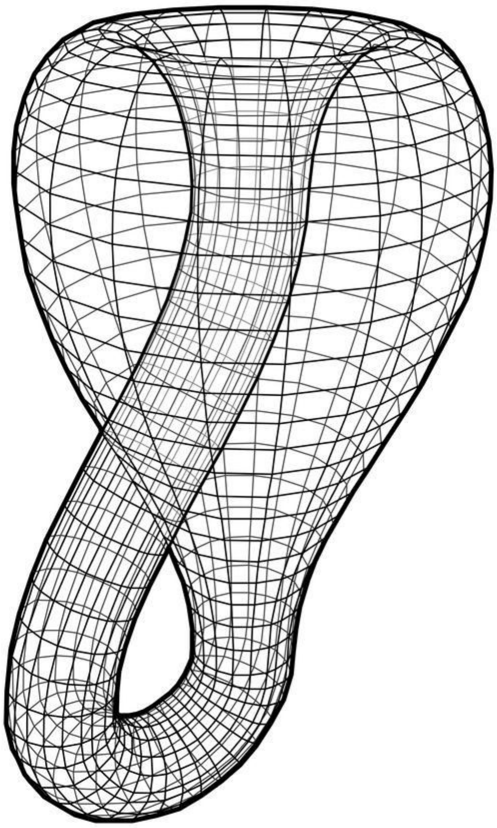
entre las fechas:

1 de abril de 2022 y 31 de julio de 2022

Firmado en la Ciudad de México, a 31 de Julio de 2022

M. en D. Pablo A. González García	**Dra. Ma. Reyna Bárcenas Hurtado**
Director del Centro	*Presidenta de Colegios Federados de Medicina de Morelos A.C.*
Representante Legal	Responsable de la Estancia
p4ablogarcia@gmail.com	barcenashurtadoreyna@gmail.com
+52 8180157253	+52 7775239828

IDC Cultura – MÉXICO
IID1407299X5
Instituto Internacional de Derecho Cultural y Desarrollo Sustentable, A.C.
C/Motozintla, 19. Col. Letrán Valle. CP. 03650. Alcaldía Benito Juárez. Ciudad de México.



Klein Bottle

(Topology as part of the Set Theory, introduced by Bertrand Russell in *Principia Mathematica*)

"Knowledge is in the end based on acknowledgement."

Ludwig Wittgenstein

Acknowledgements

This is perhaps the most difficult part of the whole doctoral thesis to write. I would not want to forget to thank anyone who has helped, accompanied or supported me on this long road to reach this point. Undoubtedly, my thesis supervisors, Prof. María Blanco, Prof. Jose Casaña and Prof. Sergio Hernández have represented an unconditional and highly valuable support for which there are not enough words to show my sincere and ample gratitude. Without them, this would not have been materialized.

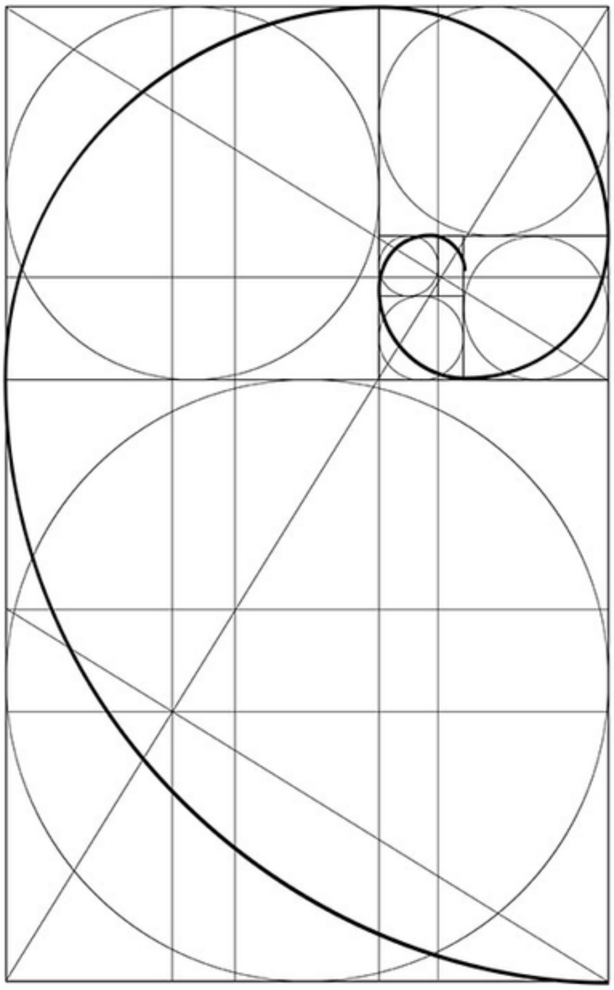
I would also like to thank the people at IDC-Cultura for their kindness and help in all aspects where their support was needed. Among them, Mr. Pablo A. González García and Dr. Ma. Reyna Bárcenas Hurtado, thank you very much for all your help. A very special thanks also to Dr. Erika Flores Deleon for her help and support for so long. There are few people as gentle and kind as she is.

To my parents for their insistence on effort and study during my first years as student. This thesis is also a consequence of that. And to my grandparents as I am sure they will be pleased to know that this has been achieved. Also, to sister and brother.

Special thanks to my in-laws, for their daily help in everything we need, which allows us to move forward in this kind of challenges we set ourselves.

To the rest of my family, due this is something they have been waiting for a long time. And to all those who have helped me in any way and whom I would like to be able to name in their entirety.

To my daughters, Gala and Alma, the greatest treasure I have in my life and who are the greatest achievement I have made together with the person I love most on the face of the earth, María. My gratitude exceeds any words I can write, any act I can perform and any thought I can have. None of it can express the love, gratitude and admiration I feel for you every day, without exception. Without you, none of this would have been possible. I love you much more than you can imagine.



Divina Proportione

"Life without love, is no life at all."

Leonardo Da Vinci

Dedicatory

To Gala

To Alma

And above all, to you, María.

$$T_H = \frac{\hbar c^3}{8 \pi K_B G M_H}$$

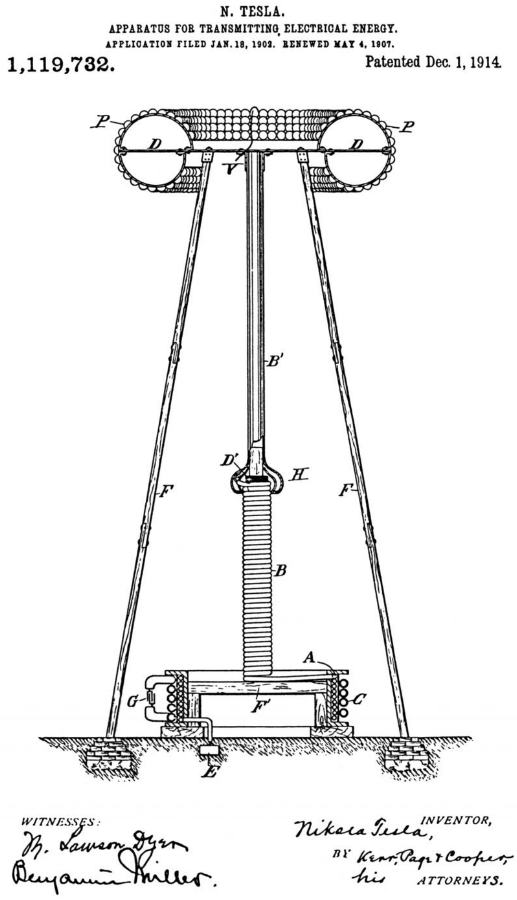
Hawking Radiation Equation

(This formula is inscribed on his gravestone, as he wished, in Westminster Abbey, between the graves of Sir Isaac Newton and Charles Darwin.)

"Science is not only a disciple of reason but also one of romance and passion."

Stephen Hawking

Index



Nikola Tesla´s patent drawing.

"The Buddhist expresses it in one way,

the Christian in another, but both say the same:

We are all one."

Nikola Tesla

Preface

The present study was carried out in the Faculty of Physiotherapy of the University of Valencia, Spain, and in the “Instituto Internacional de Derecho Cultural y Desarrollo Sustentable IDC Cultura-México”, Mexico. The supervisors of this doctoral thesis are Dra. María Blanco Díaz, Dr. Jose Casaña Granell and Dr. Sergio Hernández Sánchez.

This study received approval by the University of Valencia’s Ethics Committee of Research in Humans with the Verification Code GE3IMCG664ZM287A and has been performed in accordance with the ethical standards established in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. The ethical approval document is included in the Appendix section at the end of the main text.

No funding or grant was received for developing this study.

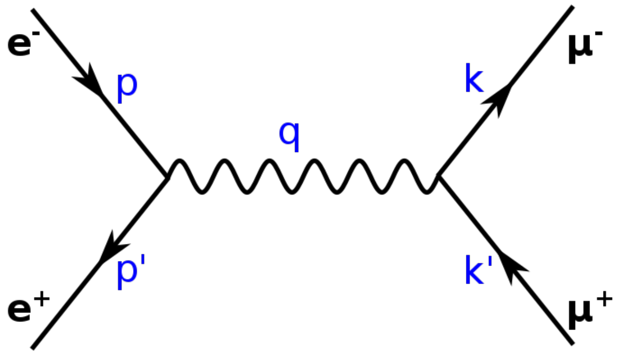
$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

Schrödinger Equation
(It governs the wave function of a quantum-mechanical system)

"Nothing is real unless it is observed"
John Gribbin

Acronyms

ACS	American Cancer Society
AI	Artificial Intelligence
APTA	American Physical Therapy Association
BC	Breast Cancer
CI	Confidence Intervals
DISCERN	Quality Criteria for Consumer Health Information Questionnaire
DL	Deep Learning
DR	Dislikes Ratio
ED	Erectile Dysfunction
FA	Factor Analysis
FR	Fisher's Ratio
GPR	Gender Parity Ratio
GQS	Global Quality Scale
H_0	Null Hypothesis
H_1	Alternative Hypothesis
HINTS	Health Information National Trends Survey
HONCode	Health On the Net Code
ICC	Intraclass Correlation Coefficient
ITU	International Telecommunication Union
KMO	Kaiser–Meyer–Olkin index
LOOCV	Leave-One-Out Cross Validation
LR	Likes Ratio
ML	Machine Learning
MSA	Measure of Sampling Adequacy
NGO	Non-Governmental Organization
NNC	Nearest Neighbor Classifier
PA	Physical Activity
PC	Prostate Cancer
PCA	Principal Component Analysis
PCs	Principal Components
PFME	Pelvic Floor Muscle Exercises
PFMT	Pelvic Floor Muscle Training
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
QoL	Quality of Life
RP	Radical Prostatectomy
SD	Standard Deviation
UI	Urinary Incontinence
VPI	Video Power Index
VR	Views Ratio
WHO	World Health Organization



Feynman Diagram

(It represents the interactions of elementary subatomic particles)

"It does not matter how many papers you have published,

if your prediction is wrong then your hypothesis is wrong. Period."

Richard P. Feynman

List Of Articles

This document comprises the doctoral thesis carried out by Álvaro Manuel Rodríguez Rodríguez, within the framework of the Doctoral Programme in Physiotherapy, code 3165, of the Universitat de València, R.D. 99/2011. The thesis is presented in the form of a compendium of publications. The articles that compose the document have been published during the years of the doctoral programme and are included in the same research line.

All the results were published in journals indexed in the Journal Citation Reports (JCR) of the Web of Science (WoS), whose impact factor, quartile and area of knowledge of the previous year to the publication date of each article are shown below.

#	Article 1	Article 2	Article 3
Title	"Quality Analysis of YouTube Videos Presenting Shoulder Exercises after Breast Cancer Surgery"	"Quality Analysis of YouTube Videos Presenting Pelvic Floor Exercises after Prostatectomy Surgery"	"Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown"
Published on	August 25th, 2021	September 15th, 2021	June 30th, 2022
DOI	10.1159/000518265	10.3390/jpm11090920	10.3390/ijerph19138016
Year of Data	2020	2020	2021
Journal	Breast Care	Journal of Personalized Medicine	International Journal of Environmental Research and Public Health
Journal Impact Factor	2.860	4.945	4.614
Category	Oncology	Medicine, General & Internal	Public, Environmental & Occupational Health
JCI Rank	181 / 310	32 / 315	45 / 182
Rank by Journal Citation Indicator (JCI)	Q 3	Q 1	Q 1
Link to JCR	https://jcr.clarivate.com/jcr-jp/journal-profile?journal=BREAST%20CARE&year=2021&fromPage=%2Fjcr%2Fhome	https://jcr.clarivate.com/jcr-jp/journal-profile?journal=J%20PERS%20MED&year=2021&fromPage=%2Fjcr%2Fhome	https://jcr.clarivate.com/jcr-jp/journal-profile?journal=INT%20J%20ENV%20RES%20PUB%20HE&year=2021&fromPage=%2Fjcr%2Fhome
Link to Article	https://www.karger.com/Article/FullText/518265	https://www.mdpi.com/2075-4426/11/9/920	https://www.mdpi.com/1660-4601/19/13/8016

The present doctoral thesis is composed of the following three studies with the following citation references:

Article 1:

Rodriguez-Rodriguez, A. M., Blanco-Diaz, M., Lopez-Diaz, P., de la Fuente-Costa, M., Duenas, L., Prieto, I. E., Calatayud, J., & Casana-Granell, J. (2021). Quality Analysis of YouTube Videos Presenting Shoulder Exercises after Breast Cancer Surgery. *Breast* Care. https://doi.org/10.1159/000518265

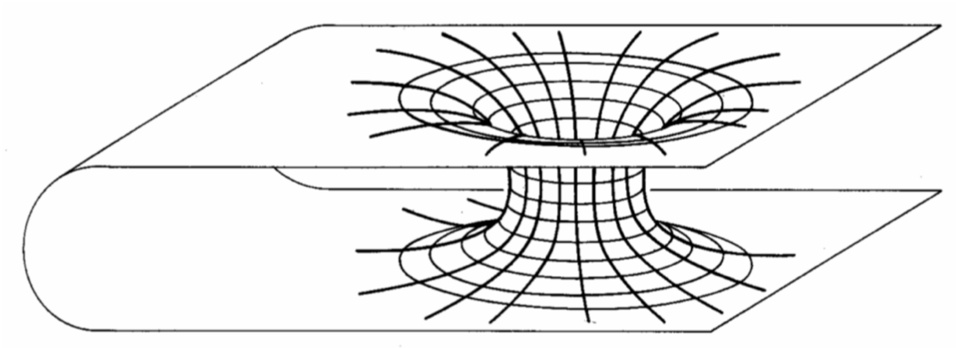
Article 2:

Rodriguez-Rodriguez, A. M., Blanco-Diaz, M., Lopez-Diaz, P., de la Fuente-Costa, M., Sousa-Fraguas, M. C., Escobio-Prieto, I., & Casaña, J. (2021). Quality analysis of youtube videos presenting pelvic floor exercises after prostatectomy surgery. *Journal of Personalized Medicine*, *11*(9), 920. https://doi.org/10.3390/jpm11090920

Article 3:

Rodriguez-Rodriguez, A. M., Blanco-Diaz, M., de la Fuente-Costa, M., Hernandez-Sanchez, S., Escobio-Prieto, I., & Casaña, J. (2022). Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown. In *International Journal of Environmental Research and Public Health* (Vol. 19, Issue 13, p. 8016). Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/ijerph19138016

Full texts of the above-mentioned studies are contained in the Appendix section, at the end of the main text of this document.



Einstein-Rosen Bridge structure connecting
disparate points in spacetime continuum.

"All thinking begins with wondering"
Socrates

Spanish Summary

Internet ha sufrido una expansión incomparable hasta convertirse en el medio más importante de difusión de información en el mundo (Kocyigit et al., 2019). Buscar información sobre salud en Internet se ha convertido en algo progresivamente común hasta el punto de que los usuarios suelen utilizar Internet como fuente primaria de información sobre temas relacionados con la salud(Baker et al., 2021; Kocyigit et al., 2019; Sui et al., 2022). Según Amante et al. (2015), casi el cincuenta por ciento de los adultos en los Estados Unidos (EE.UU.) obtienen información relacionada con la salud a través de Internet(Amante et al., 2015).

Como portal web popular para ver y compartir vídeos, YouTube es utilizada ampliamente a nivel mundial para ver y compartir vídeos (Amante et al., 2015; Baran & Yilmaz Baran, 2021; Lewis et al., 2012; Oydanich et al., 2022; Shun Zhang et al., 2020). Debido al contenido gratuito de sus vídeos y a su facilidad para llegar a la población, YouTube puede considerarse como un recurso eficaz para obtener y difundir información relacionada con la salud y utilizarse como una herramienta útil para la educación de los pacientes (Chang & Park, 2021; Jessen et al., 2022; Katz & Nandi, 2021; Warren, Wisener, et al., 2021).

Sin embargo, existen dudas razonables sobre la calidad, la fiabilidad y el contenido de los vídeos (Piskin et al., 2021; Shun Zhang et al., 2020). El sistema de carga de YouTube, permite incluir vídeos sin control ni escrutinio previos. Es por ello necesario verificar la calidad, el contenido y la exactitud de la información compartida (McMahon et al., 2022; Yildiz & Toros, 2021). Los vídeos de YouTube pueden compartir información de alta calidad relacionada con la salud pero también suscitan preocupación por los riesgos de proporcionar información de baja evidencia científica (Baran & Yilmaz Baran, 2021; Chang & Park, 2021; Culha et al., 2021; Dubey et al., 2014; Madathil et al., 2015; Oydanich et al., 2022; Patel et al., 2022; Yildiz & Toros, 2021).

Por lo tanto, el objetivo de esta tesis doctoral fue evaluar la calidad de los vídeos de YouTube en cuanto a los ejercicios recomendados que están relacionados con temas de gran importancia para la población general, como son los tipos de cáncer más comunes en las mujeres y en los hombres (cáncer de mama y de próstata, respectivamente), así como los recomendados para los períodos de confinamiento en los domicilios de la población debido a la pandemia mundial causada por el CoVid-19.

Introducción

A 30 de junio de 2022, el 69,0% de la población mundial (5.473.055.736 personas) tenía acceso a Internet. Esto significa que la tasa de crecimiento entre los años 2000 y 2022 ascendió al 1.416% (Internet World Stats, 2022). Cada año aumenta la paridad de género en el uso de Internet, según el "*Informe de Conectividad Global 2022*" publicado por la Unión Internacional de Telecomunicaciones (UIT), según el cual, a nivel mundial, el 62% de los hombres utilizaban Internet, frente al 57% de las mujeres. La brecha de género es significativamente menor en los países donde una mayor proporción de la población utiliza Internet, y es mayor en los países con un bajo uso de Internet. Además, en todas las regiones del mundo, los jóvenes (entre 15 y 24 años) son más activos en Internet que otros grupos de edad. La mayor aceptación entre los jóvenes es un buen augurio para la conectividad futura, especialmente en los países con un perfil demográfico joven o países menos desarrollados. Esto también podría mejorar las perspectivas de desarrollo de estos países, como indica dicho informe (International Telecommunication Union ITU, 2022).

Internet es un recurso valioso para la difusión de información relacionada con la promoción de la salud dirigida a la población en general (Díaz de León Castañeda & Martínez Domínguez, 2021). La disponibilidad de esta información en la web ha aumentado de forma exponencial en la última década. Un estudio realizado en 2018 mostró que los adultos de EE.UU. buscaron en la web información sobre salud un 59% más que en 2013 (Weber Shandwick AND KRC Research., 2018). El estudio también reveló que, actualmente, el 55% de las personas que buscan información sobre atención médica confían más en Internet que hace 5 años. Más del 67% de los estadounidenses buscan información relacionada con la salud en las redes sociales.

Diversos estudios han analizado el impacto positivo de los recursos cargados en las redes sociales relacionados con la salud y cómo estos pueden repercutir en la población en cuanto a la mejora de la información relativa a sus enfermedades y al conocimiento sobre el afrontamiento de estas, beneficiándose de la experiencia de personas que han padecido problemas similares (Ashtari & Taylor, 2022; Güloğlu et al., 2022; Kadakia et al., 2022; McDonough et al., 2022; Parker et al., 2021).

La reciente Encuesta de Tendencias Nacionales de Información Sanitaria (HINTS) informa de que una plataforma como YouTube tiene el potencial para compartir y difundir información oportuna relacionada con la salud, tanto en su función de repositorio de vídeos como de interfaz de red social donde los usuarios pueden interactuar y socializar(Altan Şallı et al., 2020; Madathil et al., 2015).

YouTube es el segundo motor de búsqueda más popular del mundo después de Google. Es el primer sitio web para compartir vídeos y se ha convertido en una de las plataformas principales para la difusión de información de salud (Osman et al., 2022). El 4 de abril de 2022, YouTube tenía 2.560 millones de usuarios activos mensuales y se prevé que alcance los 2.854,14 millones de usuarios en 2025. Cada minuto se transmiten más de 694 horas de vídeo y se cargan más de 500 horas cada minuto, lo que refleja el creciente apetito por el contenido de vídeo digital entre los usuarios de Internet (Statista, 2022).

Según las estadísticas publicadas, más del 95% de los internautas interactúan regularmente con YouTube a través de diversos dispositivos electrónicos, lo que facilita la accesibilidad y lo convirtie en una experiencia gratificante (Osman et al., 2022; Shameer et al., 2017; Steinhubl et al., 2015). La popularidad de YouTube, su facilidad de acceso y su carácter social lo convierten en una poderosa herramienta capaz de influir en las decisiones de las personas y promover su bienestar.

Más del 81% de los usuarios de Internet buscan información relacionada con la salud, especialmente en redes sociales, pero carecen de herramientas para evaluar los consejos recibidos (Afful-Dadzie et al., 2021; Altan Şallı et al., 2020; Ashtari & Taylor, 2022; Aydin & Akyol, 2020; Brar et al., 2022; Drozd et al., 2018; Güloğlu et al., 2022; Hamzehei et al., 2018; M. Li et al., 2022; Omnicore Agency, 2022; Osman et al., 2022; Şan, 2022; Shun Zhang et al., 2020). Así, hay que tener en cuenta la autoría, la calidad y la validez de la información contenida en los vídeos, ya que en esas plataformas no existe una selección editorial ni una evaluación de la calidad (Atci, 2019; Bahar-Ozdemir et al., 2021). Aunque muchos vídeos relacionados con la salud se consideran útiles y de alta calidad, algunos estudios revelan que no siempre es así y que algunos tienen contenido comercial diseñado para vender productos/servicios, lo que puede acarrear graves implicaciones sobre los consumidores y la toma de decisiones relacionadas con la salud al no ser precisa ni estar libre de sesgos, siendo esto un aspecto que preocupa a los proveedores de servicios sanitarios y organismos gubernamentales (Altan Şallı et al., 2020; Aydin & Akyol, 2020; Drozd et al., 2018; Hartley, 2012; Langford & Loeb, 2019; Loeb et al., 2019, 2021).

Los profesionales, instituciones y organizaciones sanitarias han de proporcionar material beneficioso para los pacientes y personas que buscan información completa y fiable en Internet(Aydin & Akyol, 2020). Los resultados de las búsquedas en YouTube se basan en la popularidad y la relevancia más que en la calidad del contenido, lo cual hace que los usuarios estén cada vez más expuestos a contenidos no verificados y potencialmente engañosos que podrían promover hábitos/actividades poco saludables (Osman et al., 2022).

Artículo 1

Según el informe Globocan 2020, el cáncer de mama es el más común en general, con una incidencia de 2.261.419 mujeres en 2020, con una mayor incidencia en los países desarrollados (The Global Cancer Observatory, 2020; Zhuang et al., 2020). Como tratamiento, las pacientes se someten a diferentes tipos de cirugía, incluyendo la mastectomía y/o la reconstrucción mamaria (American Cancer Society, 2021; Siegel et al., 2021) lo que conlleva una inmovilización prolongada causada por el miedo y/o el dolor (Reigle & Zhang, 2018). La debilidad y la limitación de la amplitud de movimiento del brazo se consideran como unas de las principales complicaciones postoperatorias (American Cancer Society, 2019, 2021) cuyos efectos perjudican seriamente su calidad de vida (Reigle & Zhang, 2018; Richmond et al., 2018).

Las mujeres que padecen cáncer de mama consultan en Youtube información detallada sobre los ejercicios recomendados (Ashtari & Taylor, 2022; Brar et al., 2021; Güloğlu et al., 2022; M. Li et al., 2022; Osman et al., 2022; Şan, 2022). Debido a que los ejercicios durante el cuidado postoperatorio mejoran la percepción del dolor y contribuyen a mejorar la calidad de vida, la Sección de Oncología de la Asociación Americana de Fisioterapia (APTA) (Worthen A., 2022), en colaboración con la Sociedad Americana del Cáncer (ACS), recomienda un programa de ejercicios postoperatorios (American Cancer Society, 2019, 2021). También Gautam et al., recomendaron un programa de ejercicios domésticos con el objetivo de mejorar los síntomas en la extremidad superior afectada, conducentes a mejorar su calidad de vida (Gautam et al., 2011; Zhou et al., 2019).

Artículo 2

Según Globocan 2020, el cáncer de próstata es el cuarto cáncer más común en general, con una incidencia de 1.414.259 en hombres, mostrando una mayor incidencia en los países desarrollados. El cáncer de próstata es una de las principales causas de enfermedad y mortalidad entre los hombres, con 375.304 fallecidos cada año (Siegel et al., 2021; The Global Cancer Observatory, 2020). La prostatectomía radical es un tratamiento común para prevenir la metástasis. Aunque la mortalidad tras la prostatectomía radical es baja (supervivencia a 5 años: 95%), la morbilidad es alta (Hodges et al., 2020). El tratamiento quirúrgico que implica la extirpación de la próstata puede provocar disfunción eréctil temporal o permanente e incontinencia urinaria, con un impacto considerable en la calidad de vida (Aydın Sayılan & Özbaş, 2018; de Lira et al., 2019; O'Callaghan et al., 2017; Pan et al., 2019; Radadiya et al., 2020). Casi el 80% de los hombres experimentan incontinencia después de la prostatectomía radical (Hodges et al., 2020).

El tratamiento de la incontinencia implica métodos terapéuticos conductuales no invasivos que consisten en la modificación de la dieta, el entrenamiento de la vejiga, los ejercicios musculares del suelo pélvico, la biorretroalimentación y la estimulación eléctrica funcional (Aydın Sayılan & Özbaş, 2018; de Lira et al., 2019; Pan et al., 2019; Prota et al., 2012).

Artículo 3

A finales de 2019, un nuevo coronavirus conocido como SARS-CoV-2 (COVID-19) surgió repentinamente en Wuhan, China (Bulut & Kato, 2020). Este virus se manifiesta como una neumonía debido a que ataca la parte inferior del tracto respiratorio en los seres humanos (Shah et al., 2020). El 31 de enero de 2020 se declaró una emergencia de salud pública internacional. Hasta el 23 de septiembre de 2022, el COVID-19 había causado más de 610.866.075 casos confirmados y más de 6.510.139 muertes tras haberse propagado por todo el mundo (World Health Organization, 2022). La mayoría de los países adoptaron políticas de confinamiento de la población en sus hogares. Los periodos prolongados de confinamiento domiciliario pueden hacer que mantenerse físicamente activo sea un gran reto (World Health Organization, 2020ª). Los confinamientos podrían ser fuentes de estrés añadido y también podrían suponer un reto para la salud mental de los ciudadanos, contribuyendo a los síntomas de ansiedad y depresión(Burtscher et al., 2020), además de aumentar otros comportamientos de riesgo para la salud. Durante los confinamientos una estrategia sanitaria proactiva debería centrarse en evitar el comportamiento sedentario (Zieff et al., 2021). La OMS define la actividad física como cualquier movimiento del cuerpo producido por sus músculos que requiere un consumo de energía (World Health Organization, 2020b). Se considera que la inactividad física constituye el cuarto factor de riesgo más importante de mortalidad (6% de las muertes en todo el mundo) y agrava la prevalencia de las enfermedades no transmisibles (World Health Organization, 2010).

La OMS recomienda en su página web, para las personas confinadas sin ninguna enfermedad respiratoria, 150 min de actividad física de intensidad moderada a la semana o 75 min de mayor intensidad, o ambas opciones combinadas. Además, la OMS recomienda "*Seguir una clase de ejercicio en línea. Aproveche la gran cantidad de clases de ejercicio en línea. Muchas de ellas son gratuitas y pueden encontrarse en YouTube. Si no tiene experiencia en la realización de* *estos ejercicios, sea prudente y consciente de sus propias limitaciones*" (World Health Organization, 2020ª).

El uso de vídeos de ejercicios orientados al fomento y la realización de la actividades físicas a través de internet, redes sociales y dispositivos digitales, son opciones factibles para mantener la salud mental y la forma física durante los períodos de reclusión (Chen et al., 2020; Gunasekeran et al., 2021; Kraus et al., 2021; Tajudeen et al., 2021). Sin embargo, no hay datos sobre la calidad de los vídeos disponibles, lo que es especialmente relevante en los periodos de confinamiento. Los vídeos subidos con contenido de salud presentan el riesgo de mostrar información engañosa/inexacta a los usuarios (Aydin & Akyol, 2020), y la autoría, la calidad y la validez de la información de los vídeos son temas a considerar (Atci, 2019; Brar et al., 2021, 2022; Güloğlu et al., 2022; M. Li et al., 2022; Osman et al., 2022; Şan, 2022).

Materiales y Métodos

Se realizaron varias búsquedas en el portal web *http://www.youtube.com/* utilizando los términos de búsqueda que se muestran en la Tabla 1. Se seleccionaron los primeros 150 vídeos disponibles para los tres artículos incluidos en este trabajo. Se pretendía replicar una estrategia de búsqueda sencilla que pudiera ser realizada por cualquier persona en YouTube. No se aplicó ninguna restricción a la búsqueda mediante filtros, por lo que YouTube ordenó los resultados por su relevancia según el algoritmo de clasificación activo en cada día concreto. Todos los vídeos se agregaron a una hoja de cálculo y se analizaron para detectar duplicidades, así como para aplicar los criterios de exclusión (Tabla 1).

Finalmente, se incluyeron 51, 133 y 68 vídeos respectivamente en cada investigación (Artículo 1 -A1-, Artículo 2 -A2- y Artículo 3 -A3-) y fueron visionados, analizados y evaluados de forma independiente por dos investigadores diferentes en cada caso durante un periodo de 5 semanas. Estos examinadores eran miembros de un grupo de investigación de sus correspondientes universidades con una larga e intensa experiencia en investigación en salud. La Fig.1 muestra las estrategias de búsqueda que se llevaron a cabo durante las diferentes fases de la revisión, según Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (Moher et al., 2009).

Medidas de resultados

Se obtuvieron las estadísticas descriptivas de los tres conjuntos de datos. Las variables que se incluyeron para este análisis fueron: la duración del vídeo, el número de visualizaciones, los días online, el ratio de visualizaciones (visualizaciones/día), el número de "Me gusta", el número de "No me gusta", el número de Suscriptores, el ratio de "Me gusta" (Nº de "Me gusta"/día), el ratio de "No me gusta" (Nº de "No me gusta"/día) y el Índice de Potencia del Vídeo (en %). Los mismos indicadores se obtuvieron para las escalas de calidad mencionadas anteriormente, que son Puntuación de la Escala Quality Criteria for Consumer Health Information Questionnaire (DISCERN), puntuación de HonCode y puntuación de GQS.

Para la evaluación de la concordancia entre examinadores, se calculó el Índice de Correlación Intraclase (ICC) en los tres artículos con un índice de confianza del 95% basado en la calificación media (k = 2), la consistencia, el modelo aleatorio de dos vías y el método de correlación de Pearson. El nivel de significación se fijó en $p<0.05$.

Según las diferentes características de los vídeos, se clasificaron en diferentes grupos atendiendo a la fuente de producción y continente de origen. Además, se recogió el número de ejercicios en comparación con el correspondiente *Gold-standard* considerado para cada artículo.

Para evaluar la popularidad de los vídeos analizados en los tres artículos, se empleó el Video Power Index (VPI) (Atci, 2019; Aydin & Akyol, 2020; Gad et al., 2022; Gun et al., 2022), que puede definirse según la siguiente fórmula:

$$VPI\ (\%) = \frac{Nº\ de\ Likes}{(Nº\ de\ Likes + Nº\ de\ Dislikes)} \times 100$$

También se evaluó la popularidad de los vídeos utilizando el View Ratio (Braczynski et al., 2021; Cakmak & Mantoglu, 2021; McMahon et al., 2022; Pamukcu & Izci Duran, 2021) que se define como sigue:

$$View\ Ratio\ (n) = \frac{Nº\ de\ Visualizaciones}{Días\ Online}$$

La calidad educativa del contenido de los vídeos seleccionados se determinó mediante la herramienta DISCERN (Bahar-Ozdemir et al., 2021; Cakmak & Mantoglu, 2021; Charnock, 1998; Di Bello et al., 2022) y el Global Quality Scale (GQS) (Kunze et al., 2019; Marwah et al., 2021; Steeb et al., 2022).

El cuestionario DISCERN fue creado por la División de Salud Pública y Atención Primaria de la Universidad de Oxford, Londres, para medir la calidad de la información relativa a los tratamientos de salud y se publicó por primera vez en 1999. Puntuaciones más altas indican mayor calidad (Azak et al., 2022; Bahar-Ozdemir et al., 2021; Charnock et al., 1999; Memioglu & Ozyasar, 2022; Oydanich et al., 2022; Shungu et al., 2021).

El GQS evalúa la calidad general del contenido de los recursos en línea. Se asigna un punto por cada uno de los 5 criterios identificables presentes en un vídeo, siendo 5 la máxima calidad educativa (Kunze et al., 2019). Esta escala cubre la accesibilidad y la calidad de la información, el flujo general de la información y la utilidad que tendría para un usuario (Azak et al., 2022; Bahar-Ozdemir et al., 2021; Duran & Kizilkan, 2021; Erdem & Karaca, 2018; Esen et al., 2019; Memioglu & Ozyasar, 2022; Pamukcu & Izci Duran, 2021; Steeb et al., 2022).

Los vídeos de los Artículos 1 y 3 también fueron evaluados con la herramienta HONCode, desarrollada por la Fundación *Health on the Net* (Boyer, 2013; *HONcode: E-Guide for Health Consumers*, 2020); una organización sin ánimo de lucro acreditada por las Naciones Unidas, que elaboró el código de conducta para estandarizar la fiabilidad de la información médica/sanitaria en línea (Boyer, 2013; Boyer et al., 1998). El HONCode es la herramienta de evaluación más utilizada para la fiabilidad y credibilidad de la información que se puede encontrar en los sitios web relacionados con la salud (Fahy et al., 2014). No pretende calificar la calidad de la información contenida en un sitio web, sino que establece una serie de reglas para que los editores de sitios web cumplan con las normas éticas básicas de suministro de información y ayuden a garantizar que los visitantes sean siempre conscientes del propósito y la fuente de los datos que están viendo (Boyer, 2013; Efe et al., 2021; Wilkens et al., 2022).

Análisis Estadístico

La Inteligencia Artificial (IA) es el estudio de problemas complejos de procesamiento de información. Su objetivo es identificar problemas útiles de procesamiento de la información y dar cuenta abstracta de cómo resolverlos. Dicha explicación se denomina "*método*" y se corresponde con un conocido teorema en matemáticas (Marr, 1977). Cada vez hay más interés por la aplicación de las técnicas de IA en la bioinformática ya que estos problemas necesitan una nueva forma de ser abordados, dada la intratabilidad de los enfoques actuales o la falta de una forma fundamentada e inteligente de explotar los datos biológicos (Hamet & Tremblay, 2017; Narayanan et al., 2002).

El aprendizaje automático (Machine Learning -ML-), cuyos métodos también se denominan algoritmos de *clustering*, es un campo dentro del área de la IA que proporciona a los sistemas la capacidad de aprender y mejorar automáticamente, basándose en la experiencia. Estos sistemas transforman los datos en información con la que aprenden a tomar decisiones. Este proceso, llevado a cabo por un algoritmo, trata de analizar y explorar los datos en busca de patrones ocultos. La aplicación de la IA en medicina tiene un importante componente virtual que está representado por el ML (también llamado Deep Learning -DL-). Existen tres tipos de algoritmos de aprendizaje automático: (i) no supervisado (capacidad de encontrar patrones), (*ii*) supervisado (algoritmos de clasificación y predicción basados en ejemplos previos) y (*iii*) aprendizaje por refuerzo (uso de secuencias de recompensas/castigos en un espacio de problemas específico) (Hamet & Tremblay, 2017; Mak & Pichika, 2019; Meskó & Görög, 2020; Panch et al., 2018). En otras palabras, el aprendizaje no supervisado se centra en descubrir la estructura subyacente o las relaciones entre las variables de un conjunto de datos y las características descubiertas a menudo pueden incorporarse a los modelos de aprendizaje supervisado (Johnson et al., 2018). Al igual que el aprendizaje supervisado, los métodos de aprendizaje automático no supervisado existen en un continuo con otros métodos estadísticos como son el Análisis de Componentes Principales, el modelado de mezclas y varios métodos de agrupación (Johnson et al., 2018). Recientemente han surgido algunas técnicas nuevas que requieren menos supuestos sobre el conjunto de datos, como los algoritmos avanzados para la factorización matricial o tensorial (Luo et al., 2017), el análisis topológico de datos (L. Li et al., 2015) y el aprendizaje profundo (Miotto et al., 2016).

El aprendizaje no supervisado difiere fundamentalmente del supervisado porque no se basa en instrucciones específicas de clasificación. En su lugar, se basa en la agrupación autónoma de objetos a través de la exploración y el descubrimiento de los patrones y estructuras subyacentes en los datos. Este enfoque se considera la forma principal en que los seres humanos recopilamos información y construimos conocimiento sobre el mundo que nos rodea. El aprendizaje no supervisado también ofrece a los humanos la capacidad de manejar cambios imprevistos en las condiciones y ampliar el conocimiento fuera de los parámetros de entrenamiento (Jha & Topol, 2016; Lecun et al., 2015; Roohi et al., 2020).

El Análisis Factorial (AF) fue introducido en 1904 por Charles Edward Spearman (Spearman, 1904) y descrito en 1995 por Bartholomew D. J. (Bartholomew, 1995). Este método permite crear nuevas variables (conocidas como factores o componentes) a partir de un conjunto de variables originales, manteniendo la máxima cantidad de información y encontrar causas ocultas (latentes) que son fuente de variabilidad en el conjunto de datos (Jain & Dubes, 1988; Jolliffe, 2002).

Suponiendo que las variables de entrada estén correlacionadas, se puede representar la misma cantidad de información con un número menor de variables. Cada variable original debe estar correlacionada con el menor número posible de factores y el número de componentes debe mantenerse al mínimo. Las saturaciones de los factores reflejan la influencia del $k^{ésimo}$ factor común en la j^{a} variable aleatoria (Bartholomew, 1995; Bartholomew et al., 2009; Ďuriš et al., 2021). Existen varios métodos que pueden utilizarse para estimar las saturaciones de los factores, conocidos como métodos de extracción de factores.

La elección de un algoritmo depende de una serie de factores que incluyen, entre otros, el tipo de datos/enfoque de aprendizaje (aprendizaje supervisado/no supervisado), la necesidad de *k*-clustering, la importancia de la precisión en el modelo elegido, la necesidad de rapidez en el análisis de los datos, los datos analizados, el tamaño del conjunto de datos, la necesidad de resultados jerárquicos y la necesidad de variables categóricas. Algunos de los métodos más comunes empleados en este enfoque incluyen la agrupación de *k-means*, la detección de anomalías o ciertos métodos estadísticos como el Análisis de Componentes Principales (PCA). Estos enfoques utilizan datos discretos o continuos como parámetro de entrada para identificar regularidades de entrada (i.e., clustering de *k-means*) y/o para reducir las representaciones dimensionales (i.e., PCA) (Celebi et al., 2013; Chu et al., 2020; Rashidi et al., 2019; Yang et al., 2022).

Mientras que las técnicas de IA y ML se utilizan ampliamente en diferentes campos de la atención sanitaria, como la definición de políticas (Bauer & Lizotte, 2021), la investigación en atención primaria (Wang et al., 2021), modelos de predicción en salud pública (Payedimarri et al., 2021), catástrofes y emergencias de salud pública (Lu et al., 2022), medios sociales y salud mental (Laacke et al., 2021), imágenes médicas (Lei et al., 2021), salud cardiovascular (Soto et al., 2021), etc., en este trabajo se utiliza el método PCA para aplicaciones de *clustering*. Otros métodos conocidos son el método de máxima verosimilitud o el de mínimos cuadrados.

Karl Pearson presentó por primera vez el PCA en 1901 (Pearson, 1901), que pretende transformar los datos de entrada multidimensionales de forma que los datos de salida se obtengan a partir de las direcciones lineales más importantes, ignorando las menos significativas. Así, se extraen las direcciones características de los datos originales y, al mismo tiempo, se reducen las dimensiones de los datos.

Este es uno de los métodos básicos de compresión de datos: las n variables originales pueden representarse con un número menor de m variables, pero explicando la mayor parte de la variabilidad del conjunto de datos original. La colección de nuevas variables (conocidas como Componentes Principales -PCs-) consiste en una combinación lineal de las variables originales. La primera PCs describe la mayor parte de la variabilidad del conjunto de datos original. Las otras PCs contribuyen a la varianza global, siempre con una proporción menor.

Todos los pares de PCs son perpendiculares entre sí (Bro & Smilde, 2014; Ďuriš et al., 2021; Gewers et al., 2021; Jolliffe, 2002). PCA y *k-means* son métodos no supervisados que utilizan datos discretos o continuos como sus parámetros de entrada para identificar regularidades de entrada (i.e., clústeres), por lo que su objetivo es agrupar un conjunto de datos médicos para caracterizar subpoblaciones de una enfermedad basada en varios parámetros clínicos de entrada conocidos y/o la reducción de la dimensionalidad del conjunto de datos (Rashidi et al., 2019; Ryu et al., 2021).

Los pasos básicos del ACP incluyen la construcción de la matriz de correlación a partir de los datos originales, el cálculo de los valores propios correspondientes, la alineación a partir de los mayores ($\lambda_1 > \lambda_2 > \cdots > \lambda_n$), el cálculo de los vectores propios de la matriz de correlación correspondientes a sus valores propios ($\upsilon_1, \upsilon_2, \ldots, \upsilon_n$), el cálculo de la variabilidad del conjunto de datos original (σ^2), la determinación del número de PCs que podrían representar las variables originales en función de la variabilidad y la transferencia de los datos originales a una nueva base de referencia según la rotación de los factores (Método de rotación ortogonal Varimax) (Bro & Smilde, 2014; Ďuriš et al., 2021; Gewers et al., 2021; Salih Hasan & Abdulazeez, 2021; Shahid et al., 2015; Yeung & Ruzzo, 2001). El PCA también puede eliminar la información redundante en las características fusionadas de grandes conjuntos de datos (Yu et al., 2022).

El número de nuevas variables puede determinarse mediante el criterio de valores propios (regla de Kaiser), cuando los factores que tienen sus valores propios $\lambda > 1$ se consideran significativos (Jolliffe, 2006b). Así, se utilizan las PCs que en conjunto explican el mayor porcentaje posible de la varianza total, o en base a una representación gráfica, el llamado gráfico Scree Plot, donde se puede ver un punto de inflexión y se deben utilizar las PCs anteriores a este punto de inflexión (Ďuriš et al., 2021; Gewers et al., 2021; Shahid et al., 2015; Yeung & Ruzzo, 2001).

Para el conjunto de datos de cada uno de los artículos se realizaron diferentes análisis estadísticos diferentes. Uno de ellos, que fue modelado sobre técnicas de aprendizaje automático de ML, considera un problema de clasificación binaria y divide mediante PCA (Jain & Dubes, 1988; Z. Zhang & Castelló, 2017) el conjunto de datos en dos grupos: "*C1*" o

"*Vídeos relevantes*" y "*C2*" o "*Vídeos no relevantes*". Así, la agrupación de los puntos del conjunto de datos se realiza utilizando algoritmos de ML no supervisado (Fisher et al., 2022; Lau et al., 2022; Rodríguez et al., 2021; Ryu et al., 2021).

La idea fundamental del PCA es mapear un espacio de características j-dimensional en un espacio i-dimensional, donde $i < j$ (Dweekat & Lam, 2022). Para este estudio, dos dimensiones (PCA1 y PCA2) fueron suficientes para agrupar los datos de cada artículo para ser representados gráficamente, con una pérdida mínima de la información del conjunto de datos original (Beam & Kohane, 2018).

Para establecer qué variables distinguen mejor los grupos en los que se dividen los vídeos en cada artículo, se utilizaron:

I. DISCERN: Los videos clasificados en el grupo "*C1/Relevant Videos*" son aquellos cuyas puntuaciones son mayores que la media; lo contrario define el segundo grupo, "*C2/Not relevant videos*".

II. Número de Ejercicios: Comparable con la variable anterior.

III. PCA: La variable binaria define ambos grupos tras un análisis de componentes principales con un algoritmo *k-means*. Con esto, los grupos *C1* y *C2* se definen por la variabilidad de la información sin pérdida de información previa. En consecuencia, en estos casos no se pierde información al reducir las dimensiones a sólo dos PCs.

Para estos análisis, se utilizaron algoritmos de agrupación para obtener el menor número de dimensiones que mejor dividiera los vídeos en dos clústeres diferentes. La precisión se estableció mediante la conocida como validación "*Leave-One-Out Cross Validation*" (LOOCV) (Vehtari et al., 2017) con el uso del clasificador "*Nearest Neighbor Classifier*" (NNC) (Kataria & Singh, 2008). El ratio de Fisher (FR) se utilizó para determinar el poder de discriminación de las variables. Cuanto mayor es el FR, más discriminatorias son, lo que es consecuencia de su baja dispersión intra-clase y su alta distancia inter-clase. En el caso de una clasificación binaria, el FR de una variable "*j*" viene determinado por:

$$\frac{(\mu_{j1} - \mu_{j2})^2}{{\sigma_{j1}}^2 + {\sigma_{j2}}^2}$$

donde μ_{ji} es la medida del centro de masas de la distribución de probabilidad de la variable "*j*" en el grupo "*i*" (i = 1, 2), mientras que σ_{ji} es la medida de la dispersión dentro de ese grupo.

Además de los FR, también se calcularon para ellos los factores de correlación de Spearman, Kendall y Pearson, que revelan la importancia, o poder de discriminación, de estas variables según los criterios de clasificación (Bonett & Wright, 2000; Croux et al., 2010; Puth et al., 2015).

De acuerdo con el PCA, Henri Kaiser (1970) presentó una medida de adecuación muestral (Kaiser, 1970, 1982).para las matrices de datos de AF que fue adaptada posteriormente por Kaiser y Rice (1974). Este factor, conocido como el índice Kaiser-Meyer-Olkin (KMO), se utilizó para determinar la normalidad multivariante y la adecuación de la muestra (Kaiser, 1982). Con el objetivo de evaluar la validez del constructo, la adecuación de la muestra para el AF se realizó mediante la prueba de esfericidad de Bartlett (Bartlett, 1951). Además, se comprobó la potencia estadística del análisis realizado mediante la combinación del análisis PCA y el índice KMO (Jackson, 1991; Park, 2021). La potencia estadística de estos dos análisis combinados indica que el análisis PCA permite clasificar los vídeos en dos grupos (*C1*, vídeos de alta calidad, y *C2*, vídeos de baja calidad) sin que se pierda información sobre ninguna de las variables consideradas en todo el estudio, lo que permite analizar la calidad de los vídeos a través de las dos PCs, en lugar de tener que hacerlo a través de todas las variables del estudio ya que su nivel de relación con las principales variables que evalúan la calidad de los vídeos es máximo y los valores *p* revelan una relación alta y estadísticamente significativa (Kaiser, 1970, 1982; Kambhatla & Leen, 1997; Salih Hasan & Abdulazeez, 2021; Yu et al., 2022; T. Zhang & Yang, 2016; X. Zhang et al., 2022; Z. Zhang & Castelló, 2017)

RESULTADOS

Tras recoger los datos de los vídeos, las listas se asignaron a dos investigadores independientes para su análisis. Se realizó un análisis ICC para los resultados de cada investigador y los datos utilizados para el análisis de cada vídeo fueron la media de las puntuaciones asignadas por cada uno de los investigadores, incluyendo las puntuaciones HONCode, GQS y DISCERN. Los datos básicos de las estadísticas descriptivas de los resultados medios de VPI, HONCode, GQS y DISCERN se muestran en la Tabla 3. Con el mencionado intervalo de confianza del 95%, y el nivel de significación fijado en $p<0{,}05$, los valores resultantes se clasifican en la Tabla 2 (Koo & Li, 2016). La concordancia entre investigadores, para los artículos incluidos en este estudio, obtuvo una puntuación de 0.9860 en el Artículo 1, de 0.9505 en el Artículo 2 y de 0.9299 en el Artículo 3.

Para el Artículo 1 se utilizó la versión ampliada de la escala DISCERN, mientras que para los Artículos 2 y 3 se utilizó la versión reducida. Con el fin de facilitar la comparación de estos resultados, se aplicó un cambio de escala a los resultados del Artículo 1, pasando de una escala de 0-80 a una escala de 0-5. Los resultados se muestran en la Tabla 5. En esta tabla, se puede observar que, en el Artículo 1 y en el Artículo 3, las Instituciones Académicas obtienen la puntuación más alta (3.52 y 3.25, respectivamente) y las Instituciones Sanitarias obtienen la segunda puntuación más alta (3.43 y 2.83, respectivamente). En el Artículo 2, estos dos tipos de instituciones obtienen la misma puntuación (3.88, ambos) por detrás de las ONGs que obtienen la puntuación más alta (4.33). Los medios de comunicación y las personas no vinculadas al ámbito sanitario obtienen las puntuaciones más bajas en los tres artículos y, sólo en el tercero, las ONGs se incluyen también en el grupo de resultados más bajos.

En cuanto al continente origen de los vídeos, en el Artículo 1 la puntuación más alta la obtienen los vídeos producidos en América (3.23), seguidos de cerca por los producidos en Europa (3.22). En el Artículo 2 y en el Artículo 3, los vídeos producidos en Australia obtienen la puntuación más alta (3.55 y 3.50, respectivamente), seguidos por los vídeos producidos en Europa (3.40 y 2.58, respectivamente). En los tres artículos, los vídeos producidos en África obtienen la puntuación más baja de todos (0.00, 0.00 y 2.00, respectivamente), al igual que con las puntuaciones de GQS.

La Tabla 6 muestra los resultados medios de las puntuaciones GQS según la fuente de producción (autor) y el continente de origen en cada uno de los artículos de esta investigación. En dicha Tabla 6 se puede observar que las Instituciones Académicas, las Instituciones Sanitarias y las ONGs obtienen las puntuaciones más altas (en diferente orden, pero siempre estos tres tipos de autorías) en la escala GQS. Las Instituciones Académicas obtienen la mayor puntuación en el Artículo 1 (3.83) y en el Artículo 2 (4.00, mismo resultado que las ONGs), seguidas en ambos casos por las Instituciones Sanitarias (3.74 en el Artículo 1 y 3.77 en el Artículo 2). En el Artículo 3, las Instituciones Sanitarias obtienen la puntuación más alta (3.33), seguidas por las ONGs y las Instituciones Académicas, que obtienen puntuaciones cercanas (3.06 y 3.00, respectivamente). En los tres artículos, las puntuaciones más bajas las obtienen las personas no vinculadas al ámbito sanitario (3.16, 2.82 y 1.84, respectivamente), seguidos de los Medios de Comunicación (3.31, 2.96 y 2.42, respectivamente).

En cuanto al origen de los vídeos, en el Artículo 1 los vídeos producidos en Europa recibieron las puntuaciones más altas (3.79), seguidos de los producidos en América (3.52). Como ocurrió en la escala DISCERN, en el Artículo 2 y en el Artículo 3, los vídeos producidos en Australia recibieron las puntuaciones más altas (3.55 y 3.50, respectivamente), seguidos de los producidos en Europa (3.40 y 2.54, respectivamente).

El nivel de calidad de los vídeos puede clasificarse como se muestra en la Tabla 7 tanto para los valores de DISCERN como de GQS (Charnock et al., 1999; Erdem & Karaca, 2018; Esen et al., 2019; Kunze et al., 2019; Yalkin et al., 2022).

En cuanto a los resultados del HONCode, en el Artículo 1, las Instituciones Académicas recibieron la puntuación más alta (62.96), seguidas de las ONG (28.78) y los Medios de Comunicación la más baja (10.40). En el Artículo 3, las Instituciones Sanitarias recibieron la puntuación más alta (72.33), seguidas de las ONG (67.43) y las Instituciones Académicas (66.63) y las personas no vinculadas al ámbito sanitario la puntuación más baja (53.06). Según su origen, en el Artículo 1 los vídeos con mejor puntuación son los producidos en Asia (50.69), seguidos de los vídeos de Europa (45.03). En el Artículo 3, los vídeos de Australia recibieron la puntuación más alta (69.00). Como ya se ha explicado, el instrumento HONCode no estaba disponible para el Artículo 2, razón por la cual no hay resultados de dicha herramienta para este segundo artículo. Todos estos resultados se muestran en la Tabla 8.

Los coeficientes de correlación de Pearson para las variables GQS y HONCode respecto de la variable DISCERN se muestran en la Tabla 9.

El índice KMO, para la adecuación de la muestra, se realizó para comprobar la conformidad del tamaño de la muestra antes de aplicar el AF (Pechenizkiy et al., 2004). El resultado de esta prueba mostró que el valor KMO alcanzó 0.916 en el Artículo 1, 0.872 en el Artículo 2 y 0.845 en el Artículo 3, datos que se muestran en la Tabla 11, en la que también se muestran las varianzas explicadas totales del PCA para cada artículo. Según esto, se puede determinar que los tamaños de las muestras eran, según Kaiser y su indicador KMO, "*Meritorios*" para estos conjuntos de datos. Se han representado también los Scree Plots que muestran la variación explicada para cada PC en orden descendente (Figura 2, Figura 4 y Figura 6), que se utiliza en el análisis PCA para elegir el número de PCs (Abdi & Williams, 2010; Ďuriš et al., 2021; Eastment & Krzanowski, 1982; Gewers et al., 2021; Hasan & Tahir, 2010; Horn, 1965; Ivosev et al., 2008; Kanyongo, 2005; Ledesma et al., 2015; Qu et al., 2002; Shahid et al., 2015; Silva et al., 2020; Yeung & Ruzzo, 2001; Z. Zhang & Castelló, 2017; Zhu & Ghodsi, 2006).

Representado gráficamente cada vídeo por un punto en un sistema de ejes coordenados donde cada eje es una de las PCs (Gewers et al., 2021; Jolliffe, 2006ª; Mahmoudi et al., 2021; Saerens et al., 2004; Yu et al., 2022) se permite una clara visualización de la autoorganización a la que se someten los vídeos sin ninguna intervención manual. La clara diferenciación de las nubes de puntos en las que se agrupan los vídeos de alta calidad, por un lado, y los de baja calidad, por otro, muestran claramente que esta agrupación natural permite una visualización intuitiva de los niveles de calidad de cada una de las nubes de puntos.

DISCUSIÓN

Esta tesis se realizó con el fin de comprender mejor la naturaleza de la información a la que acceden los pacientes en una plataforma masiva en línea cuyo uso es cada vez mayor. El creciente uso de las nuevas tecnologías y la gran cantidad de contenidos disponibles han convertido a Internet en una importante fuente de información sanitaria tanto para los profesionales de la salud (que las utilizan con mayor frecuencia para comunicarse con los pacientes y lograr la necesaria adherencia e interacción en los procesos sanitarios) como para los pacientes (que las utilizan como fuente de información). Los vídeos son una fuente de información sencilla, atractiva y asequible, con una gran capacidad de difusión disponible, pero que no está sujeta a ningún tipo de inspección. En consecuencia, esto puede llevar a que se transmitan mensajes erróneos e incluso perjudiciales para los usuarios.

En el caso de los vídeos analizados, la mayor parte de ellos fueron producidos por Instituciones Sanitarias, lo que puede considerarse como una garantía, debido a que se encuentran en los primeros lugares de las escalas de calidad. En el caso del Artículo 3, los vídeos producidos por Instituciones Sanitarias representaron un pequeño porcentaje del total, pero también alcanzaron puntuaciones muy altas en el análisis de calidad, lo que puede considerarse una garantía de la información compartida en esos vídeos.

La concordancia inter-investigadores para estos estudios fue calculada mediante el ICC (Koo & Li, 2016) y puntuó, respectivamente para los tres artículos, 0.9860, 0.9505 y 0.9299, apuntando hacia una concordancia "*Excelente*", según los intervalos del ICC mostrados en la Tabla 2.

La duración media de los vídeos evaluados en esta tesis doctoral fue de 9:42 min para el Artículo 1, 14:42 min para el Artículo 2 y 10:37 min para el Artículo 3, lo que se alinea con estudios anteriores que informaron de una duración media de 6:17-10:35 min (Gokcen & Gumussuyu, 2019; Kuru & Erken, 2020), excepto en el Artículo 2 ya que tiene la mayor desviación estándar y además el 63.91% de los vídeos tienen una duración inferior a 10:00 min, lo que señala una alta similitud con estudios anteriores. No se encontró ninguna relación estadísticamente significativa entre la duración de los vídeos y su calidad o popularidad, al contrario que en el estudio de Akif Aydin, que encontraron que los vídeos con mayor duración y mayor VPI parecían estar asociados con puntuaciones de mayor calidad (Aydin & Akyol, 2020).

Aunque el número de ejercicios recomendados podría ser mayor, las puntuaciones de los vídeos son "*Alta*" y "*Media*" en las escalas de calidad que se consideraron para los artículos, lo que coincide con otros autores (Akyol & Karahan, 2020; Altan Şallı et al., 2020; Karakoyun & Yildirim, 2021; Meteran et al., 2022; Morra et al., 2022; Nangia et al., 2020; Navarro et al., 2021; Onder & Zengin, 2021; Salman & Bayar, 2021). En cuanto al número de ejercicios recomendados analizados en cada uno de los artículos, se obtuvieron los siguientes resultados:

Artículo 1

La Sección de Oncología de la Asociación Americana de Fisioterapia (APTA), en colaboración con la Sociedad Americana del Cáncer (ACS), recomienda un programa de siete ejercicios. De estos siete, los vídeos analizados muestran una media de 1.75 ejercicios (lo que representa el 25% de los ejercicios recomendados). Los vídeos que obtuvieron una puntuación de Alta calidad muestran un 30% de los ejercicios recomendados (media de 2.10 ejercicios).

Artículo 2

El número medio de ejercicios recomendados, según las normas estandarizadas, es de 1.93 ejercicios. Los vídeos de muy alta calidad muestran una media de 3.58 ejercicios.

Artículo 3

De los 12 ejercicios sugeridos por la OMS, los vídeos disponibles en YouTube relativos a la actividad física durante el confinamiento causado por el CoVid-19 sólo muestran una media de 1.72 ejercicios, 2.83 en los vídeos de alta calidad, no habiendo ningún vídeo que haya mostrado más de 5 ejercicios.

CONCLUSIONES

De la información obtenida en la recogida de datos y de los diferentes análisis estadísticos realizados para estos tres estudios que componen el objeto de investigación de esta tesis, se pueden extraer las siguientes conclusiones:

Artículo 1

La calidad de la información ofrecida en YouTube sobre los ejercicios postoperatorios de hombro recomendados para pacientes con cáncer de mama, según las puntuaciones obtenidas en las escalas de calidad, es "*Alta*".

Artículo 2

La calidad de los vídeos en YouTube sobre los ejercicios recomendados para el suelo pélvico de los pacientes con cáncer de próstata tras una prostatectomía, según las puntuaciones obtenidas en las escalas de calidad, es "*Alta*".

Artículo 3

La calidad de los vídeos disponibles en YouTube relativos a la actividad física durante el bloqueo de CoVid-19 es ligeramente inferior a la de los dos artículos anteriores y no refleja las recomendaciones de la OMS. No obstante, según las escalas de calidad, la calidad de estos vídeos es "*Media*".

LIMITACIONES Y PERSPECTIVAS FUTURAS

Limitaciones

Con el fin de simular el comportamiento estándar de los usuarios, no se realizaron búsquedas en modo incógnito para evitar la influencia del historial de navegación y la ubicación geográfica. Aunque el contenido de YouTube cambia con el tiempo, estos análisis representan el estado de los vídeos en un momento concreto. Sólo se incluyeron en los artículos los vídeos a los que se accedió directamente en YouTube al utilizar los términos de búsqueda indicados en la Tabla 1. En los análisis realizados se excluyeron los enlaces externos a otros sitios web.

Debido a que esta tesis doctoral trató de reproducir el comportamiento de búsqueda típico del usuario medio, las búsquedas se limitaron a los primeros 150 vídeos. Como otros estudios han explicado anteriormente, la mayoría de los internautas no prestan atención más allá de los primeros 50 resultados de búsqueda.

Perspectivas Futuras

Se considera necesario un proceso de verificación/validación de la información disponible en la web, así como programas educativos para facilitar el acceso de las personas a la información más fiable (Ahmad et al., 2020; Bahar-Ozdemir et al., 2021; Fode et al., 2020; Warren, Sawhney, et al., 2021). Las instituciones educativas y sanitarias, así como los profesionales de la salud, las autoridades sanitarias gubernamentales y los responsables políticos deben participar en el desarrollo de políticas que mejoren la información disponible en la web con el objetivo de crear un impacto positivo en el comportamiento de las personas relacionado con la salud, como se recomienda en otros estudios (Bahar-Ozdemir et al., 2021; Brar et al., 2021; Gulve et al., 2022; Kodonas & Fardi, 2021; Warren, Sawhney, et al., 2021). Se necesitan estrategias y políticas eficaces capaces de indicar la calidad de esta información para filtrar la errónea o no rigurosa que pueda afectar a la salud de las personas. Estas herramientas deberían ayudar a cualquier usuario/espectador a distinguir los vídeos de alta y baja calidad (Enver et al., 2020; Fode et al., 2020; Yalkin et al., 2022; X. Zhang et al., 2022). Dado que no existe una marca para identificar los vídeos de alta calidad, se recomienda para su fácil identificación que los pacientes atiendan a la fuente de producción, ya que las instituciones académicas, las ONG y las instituciones sanitarias muestran cifras más altas en las escalas de calidad. Los vídeos de más alta calidad fueron subidos por fuentes académicas y profesionales de la salud (Fode et al., 2020; Ozduran & Büyükçoban, 2022), por lo que estos grupos deberían promover más vídeos de alta calidad (Crutchfield et al., 2021; Goobie et al., 2019; Güloğlu et al., 2022; Lang et al., 2022; Morra et al., 2022; Passos et al., 2020). También sería recomendable descartar aquellos vídeos que contengan anuncios publicitarios y puedan presentar algún conflicto de intereses (Brar et al., 2021; Vu et al., 2021).

Con el apoyo de técnicas de aprendizaje automático y algoritmos de IA no supervisados, la informática puede hacer avanzar la medicina basada en la evidencia a través de herramientas de apoyo (Amezcua-Prieto et al., 2020; Bauer & Lizotte, 2021; Yoldemir, 2020). Esta es una gran oportunidad para la transformación del sistema sanitario, ya que el coste de aumentar la capacidad de toma de decisiones en todo el sistema sanitario no es muy elevado (Benke & Benke, 2018; Panch et al., 2018; Schwalbe & Wahl, 2020; Singh, 2019; Thiébaut & Thiessard, 2018).

Son necesarios estudios e intervenciones que impliquen a las instituciones, guiados por los principios de la investigación médica traslacional, para aumentar la calidad de la información sanitaria en los vídeos de YouTube. Mediante el desarrollo de políticas eficaces que mejoren las conductas y comportamientos relacionados con la salud de la población, se debe de conseguir diferenciar la información fiable y de buena calidad, evitando los riesgos inherentes en las búsquedas de información relacionada con la salud en las redes sociales.

$$(\partial + m)\psi = 0$$

Dirac Equation combining the principles of quantum mechanics and
the theory of special relativity implying the antimatter.
Referred to as he most beautiful equation known in physics which, due the
phenomenon of quantum entanglement, is known as the "*Equation of Love*".

"If you are receptive and humble,
mathematics will lead you by the hand"
Paul Dirac

1

1. Abstract

The Internet has been expanded worldwide to become the most important mean to spread information in the world (Kocyigit et al., 2019). Searching and finding health information using the Internet has become progressively common while people often use the Internet as a source of health information (Baker et al., 2021; Kocyigit et al., 2019; Sui et al., 2022). According to Amante et al (2015), nearly fifty per cent of adults in the United States (US) get health-related information on the Internet (Amante et al., 2015). As a popular video sharing site, YouTube is extensively used all around the world for users to watch and share videos (Amante et al., 2015; Baran & Yilmaz Baran, 2021; Lewis et al., 2012; Oydanich et al., 2022; Shun Zhang et al., 2020). Due to the free content of its videos and its ease of reach to the population, YouTube can be considered an effective resource for obtaining and disseminating health-related information. Consequently, it can also be used as a useful tool for patient education (Chang & Park, 2021; Jessen et al., 2022; Katz & Nandi, 2021; Warren, Wisener, et al., 2021). However, there are doubts about the quality, reliability and content of the videos (Piskin et al., 2021; Shun Zhang et al., 2020). Especially in view of YouTube's philosophy, where anyone can upload videos without prior check or scrutiny, and which could be used for promotional ends, it is necessary to verify the quality, content and accuracy of the information shared in the uploaded videos (McMahon et al., 2022; Yildiz & Toros, 2021). This means that YouTube videos may raise concerns about the risks of providing misleading health-related information in videos available on this platform (Baran & Yilmaz Baran, 2021; Culha et al., 2021; Dubey et al., 2014; Madathil et al., 2015; Patel et al., 2022).

A previous systematic review investigating eighteen pieces of research found that YouTube can display misleading and conflicting health-related information, but, on the other hand, it can also share high quality health-related information (Chang & Park, 2021; Oydanich et al., 2022; Yildiz & Toros, 2021). So, digital platforms are a promising way to support physical activity levels and may have provided an alternative for people to maintain their activity while at home (Güloğlu et al., 2022; Kadakia et al., 2022; McDonough et al., 2022; Parker et al., 2021).

Therefore, the aim of this study was to assess the quality of YouTube videos, that any internet user could access, regarding to the recommended exercises that are related to topics of major importance for general population. In this regard, exercises related to the most common types of cancer in women and men (breast cancer and prostate cancer, respectively) have been considered, as well as those recommended for periods of confinement due to the global pandemic caused by CoVid-19.

1.1. Article 1

The prolonged immobilization suggested after breast cancer (BC) surgery causes morbidity. Patients search the Internet, especially social networks, for recommended exercises. The aim of this observational study was to assess the quality of YouTube videos, accessible for any patient, about exercises after BC surgery (Rodriguez-Rodriguez, Blanco-Diaz, Lopez-Diaz, de la Fuente-Costa, Duenas, et al., 2021).

1.2. Article 2

Prostate cancer (PC) is a major cause of disease and mortality among men. Surgical treatment involving the removal of the prostate may result in temporary or permanent erectile dysfunction (ED) and urinary incontinence (UI), with considerable impact on quality of life. (QoL) Pelvic floor muscle training (PFMT) is one of the recommended techniques for the prevention, treatment, and rehabilitation of postoperative complications. The aim of this observational study was to assess the quality of YouTube videos related to exercises after prostatectomy surgery (Rodriguez-Rodriguez, Blanco-Diaz, Lopez-Diaz, de la Fuente-Costa, Sousa-Fraguas, et al., 2021).

1.3. Article 3

The world has been experiencing a pandemic caused by COVID-19. Insufficient physical activity can increase the risk of illness. The aim of this study was to evaluate the quality of YouTube videos related to home exercises during lockdown and their adherence to WHO recommendations after replicating a simple search process that could be performed by any individual internet user (Rodriguez-Rodriguez et al., 2022).

Möbius strip

(Single-sided and single-edged surface that has the mathematical property of being a non-orientable object)

"Everything has beauty,

but not everyone sees it"

Confucius

2

2. Introduction

As of June 30th, 2022, an estimated 69.0% of the global population, i.e., 5,473,055,736 people, had access to the Internet. These figures mean that the growth rate between year 2000 and 2022 goes up to 1,416% (Internet World Stats, 2022). Every year the gender parity is growing up in the usage of internet and, according to the "Global Connectivity Report 2022" published by the International Telecommunication Union (ITU), globally, 62 per cent of all men were using the Internet, compared with 57 per cent of all women. The gender gap is significantly smaller in countries where a higher proportion of the population uses the Internet, and a higher gender gap exists in countries with low Internet use. In countries where everyone is using the Internet, by definition, there must be gender parity. The gender parity ratio (GPR) is calculated as the proportion of women using the Internet divided by the proportion of men using the Internet. A value smaller than 1 indicates a larger proportion among men than among women. A value greater than 1 indicates the opposite. Values between 0.98 and 1.02

reflect gender parity as established in the 2030 targets. Also, in all regions of the world, young people (between 15 and 24 years old) are more active on the Internet than other age groups. Data published in the mentioned report shows the huge divide when comparing the Internet uptake of people of 75 years old and above, and of those between 15 and 24. Greater uptake among the young bodes well for future connectivity, particularly in countries with a young demographic profile. In least developed countries (LDCs) for example, where half the population is less than 20 years old, the workforce will become more connected and digitally skilled as this young generation joins its ranks. This in turn will improve the development prospects of these countries, as stated in such report (International Telecommunication Union ITU, 2022).

The Internet has the potential to be a valuable resource for the dissemination of health promotion information to the general population, mainly in conditions with well-developed health and digital literacy (Díaz de León Castañeda & Martínez Domínguez, 2021). The availability of health information on the web has increased dramatically over the past decade. A study in 2018 showed that adults in the United States looked on the web for health information 59% more than in 2013 (Weber Shandwick AND KRC Research., 2018). The study also revealed that, currently, 55% of health care information seekers are relying more on the internet and web-based resources for their health-related information than 5 years ago. More than 67% of American health care information seekers mentioned that they look for health information on social media. The importance of online resources uploaded on social platforms has been studied among different people with different diseases and different studies have discussed their positive impact and how patients can benefit from the emotional and community support. People who faced similar problems can share possible solutions, suggest how to cope with symptoms, and provide information about their disease to other members of the group (Ashtari & Taylor, 2022).

YouTube is one of the most popular internet sites, in fact, YouTube is the world's second most popular search engine (after Google) and social media platform, it is the first video-sharing website and also operates as a primary Internet platform for consumer-targeted health information (Osman et al., 2022). As of April 4th, 2022, YouTube had 2.56 billion active users monthly and it is projected to reach 2,854.14 million users by 2025. More than 694 hours of video were streamed on YouTube every minute and more than 500 hours of video were uploaded to YouTube every minute, reflecting the increased appetite for digital video content among internet users (Statista, 2022).

According to published statistics, over 95% of the Internet population are regularly interacting with YouTube via personal computers, laptops, tablets, or mobile phones which makes the YouTube video experience much more enjoyable and available to all users on-demand, anywhere and anytime (Osman et al., 2022; Shameer et al., 2017; Steinhubl et al., 2015). YouTube's popularity, ease of access, and social nature made it a powerful tool for influencing individuals' decisions and promoting their well-being.

YouTube is the leading and the largest video hosting platform which has been typically utilized as a health information resource and medical education tool by Internet users (Bahar-Ozdemir et al., 2021). A recent Health Information National Trends Survey (HINTS) reports a substantial increase in the use of the Internet for retrieving health information. Recent surveys have found that 8 of 10 Internet users accessed health information online. These results suggest that a platform like YouTube has the potential to serve as an important vehicle for sharing and disseminating timely health-related information, both in its function as a repository of videos and as a social networking interface where users can interact and socialize (Altan Şallı et al., 2020; Madathil et al., 2015). Mamlin et al (2016) anticipated that social media websites, such as YouTube, would be widely used for (1.) exchanging healthcare related information between healthcare providers and consumers, (2.) facilitating peer-to-peer support for patients, and (3.) enhancing public health surveillance (Mamlin & Tierney, 2016).

Advances in eHealth technology have cultivated transactional opportunities for patients to access, share, and monitor health information (Paige et al., 2018). Patients and health professionals often search on Internet for information regarding many health-related topics. It has been proven that 81% of all Internet users search for information related with health (Omnicore Agency, 2022). Generally, patients search for detailed information about recommended exercises related with different pathologies. Therefore, the authorship, quality, and validity of the information contained in the videos must be considered (Atci, 2019). So, many patients search for medical information on internet and, also in the social media platforms, but they lack the tools to evaluate the advice given (Drozd et al., 2018).

Although many health-related videos on YouTube are deemed educationally useful and of high quality, some studies reveal that this is not always the case and some have commercial content designed to sell products or services, which may have serious implications for consumer attitudes and medical decision making (Langford & Loeb, 2019; Loeb et al., 2019, 2021).

Nowadays, the fact that people depend on Internet for health education has increased dramatically. More than 70% of the United States adult population is using Internet as the first source for health-related information, which leads to the appearance of the so-called e-patient, where the Internet served as a source to obtain information that the patient failed to obtain from his/her healthcare providers during their clinic appointment (Alefishat et al., 2021). From these websites, patients may have access to a large amount of health-related information which is available for free to them (Hamzehei et al., 2018). As health information seeking on the Internet has become more popular, the number of patients using the online platforms as a source of medical information about diseases and treatment methods has increased accordingly (Bahar-Ozdemir et al., 2021). This means that video-sharing websites must be understood as social media sites where there is no editorial selection or quality assessment. Likewise, potentially harmful and inaccurate information about science and biomedical topics can be disseminated. The kind of content stored on YouTube and the quality of the information are unclear (Hartley, 2012). Social media has great potential to provide easy access to medical information, but it is likely that the information received is neither accurate nor free of bias (Altan Şallı et al., 2020; Drozd et al., 2018).

One question that has been highlighted by biomedical institutions is whether YouTube provides users and (potential) patients with accurate and helpful information or if the videos are possibly harmful and misleading. Healthcare professionals and organizations should be encouraged to provide more beneficial material and animated videos to people looking for comprehensive, reliable information on Internet (Aydin & Akyol, 2020). Healthcare providers and government agencies have expressed their concern about the veracity and quality of the information available on this platform, as the health-content videos uploaded by various sources can be misleading and present inaccurate information to patients (Aydin & Akyol, 2020). YouTube's search results are based on popularity and relevancy rather than content quality. This creates an issue for informal or unguided learners who are increasingly exposed to unverified and partly misleading content that could promote unhealthy habits and activities (Osman et al., 2022).

2.1. Article 1

According to Globocan 2020 report, breast cancer (BC) is the most common cancer overall with an incidence of 2,261,419 women in 2020 and exhibits higher incidence in developed countries. There were over 2 million new cases in 2020 (The Global Cancer Observatory, 2020; Zhuang et al., 2020). Women with BC are often treated with different types of surgery, including mastectomy and/or breast reconstruction (American Cancer Society, 2021; Siegel et al., 2021) which leads to a prolonged immobilization caused by fear and/or pain (Reigle & Zhang, 2018). The weakness and limitation of patients' range of motion of the arm caused by pain and stiffness are considered one of the major postoperative complications of BC treatment (American Cancer Society, 2019, 2021) whose effects impairs their QoL (Reigle & Zhang, 2018; Richmond et al., 2018).

Exercises during postoperative care of women with BC are important because it improves pain perception and contributes to enhancing QoL. The Oncology Section of the American Physical Therapy Association (APTA) (Worthen A., 2022), in collaboration with the American Cancer Society (ACS), recommends a series of seven exercises (American Cancer Society, 2019, 2021). A home-based exercise program improved symptoms in the affected upper limb and led to improved QoL in women (Gautam et al., 2011; Zhou et al., 2019). Generally, patients search on Youtube for detailed information about recommended exercises for improving their QoL (Ashtari & Taylor, 2022; Brar et al., 2021; Güloğlu et al., 2022; M. Li et al., 2022; Osman et al., 2022; Şan, 2022).

2.2. Article 2

According to Globocan 2020, prostate cancer is the fourth most common cancer overall, with an incidence of 1,414,259 in men in 2020, exhibiting higher incidence in developed countries. Prostate cancer is a major cause of disease and mortality among men, with 375,304 men dying of it each year (Siegel et al., 2021; The Global Cancer Observatory, 2020). Radical prostatectomy (RP) is a common curative treatment to prevent metastasis. Although mortality after RP is low (5-year survival: 95%), morbidity is high (Hodges et al., 2020). Surgical treatment involving the removal of the prostate may result in temporary or permanent erectile dysfunction (ED) and urinary incontinence (UI), with considerable impact on QoL (QoL) (Aydın Sayılan & Özbaş, 2018; de Lira et al., 2019; O'Callaghan et al., 2017; Pan et al., 2019; Radadiya et al., 2020). Depending on how continence is defined, almost 80% of men experience incontinence after RP (Hodges et al., 2020).

Treatment of incontinence involves noninvasive behavioral therapeutic methods consisting of diet modification, bladder training, pelvic floor muscle exercises (PFME), biofeedback, and functional electrical stimulation (Aydın Sayılan & Özbaş, 2018). Pelvic floor exercises have been used to improve urinary continence following RP, with good results (de Lira et al., 2019; Prota et al., 2012). Urinary continence can be achieved through contraction training of the pelvic floor muscles (Pan et al., 2019). PFMT is one of the recommended techniques for the prevention, treatment, and rehabilitation of RP-related complications (de Lira et al., 2019). It can improve UI and ED after prostatectomy (Pan et al., 2019). Many patients search the Internet for medical information, but they lack the tools to evaluate the advice provided (Afful-Dadzie et al., 2021; Altan Şallı et al., 2020; Ashtari & Taylor, 2022; Aydin & Akyol, 2020; Brar et al., 2022; Drozd et al., 2018; Güloğlu et al., 2022; M. Li et al., 2022; Osman et al., 2022; Şan, 2022; Shun Zhang et al., 2020).

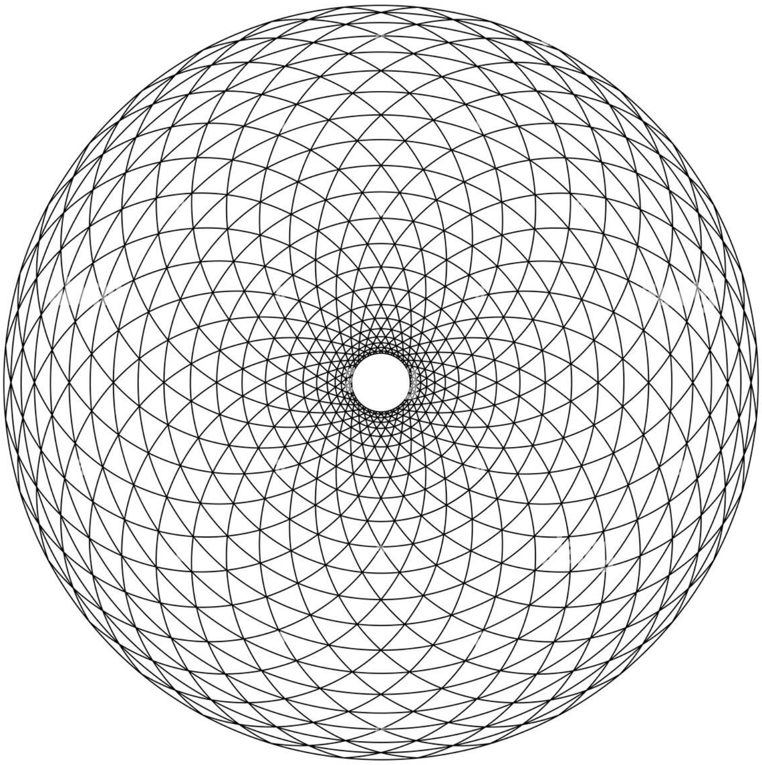
2.3. Article 3

By the end of 2019, a novel coronavirus known as SARS-CoV-2 (COVID-19) suddenly arose in Wuhan, China (Bulut & Kato, 2020). This virus manifests as pneumonia due to the fact that it attacks the lower part of the respiratory tract in humans (Shah et al., 2020). An international public health emergency was declared on 31 January 2020. As of September 23th 2022, COVID-19 caused over 610,866,075 confirmed cumulative cases and over 6,510,139 cumulative deaths after being spread worldwide (World Health Organization, 2022). Most countries have adopted mandatory home lockdown policies. However, prolonged periods of time at home can make staying physically active a major challenge (World Health Organization, 2020a). WHO defines physical activity (PA) as any movement of the body produced by its muscles that requires energy consumption, including exercise and other activities that include physical movement and are conducted as part of play, work, active transport, household chores, and leisure activities (World Health Organization, 2020b). Self-quarantine and prolonged stays at home could be sources of added stress and could also pose challenges for citizens' mental health, contributing to anxiety and depression symptoms (Burtscher et al., 2020), as well as increase other health risk behaviours. During the COVID-19 lockdown (Zieff et al., 2021), a proactive health strategy should be focused on avoiding sedentary behaviour. In addition, insufficient PA constitutes the fourth most important risk factor for mortality (6% of deaths worldwide) and severely influences the prevalence of non-communicable diseases (World Health Organization, 2010).

Very few guidelines targeting public health in general consider PA routines on a daily basis for people that live in different levels of isolation during the COVID-19 pandemic (Matias et al., 2020). Home exercise, through simple, safe, and easily implementable exercises, is well recommended to maintain good fitness levels. These exercises should include, but not be limited to, activities for strengthening, stretching, and improving balance and control and/or a combination of them (Chen et al., 2020).

WHO recommends, in a guide published on its website, for confined people without any respiratory illness or suspect of it, 150 min a week of moderated-intensity PA or 75 min a week of higher-intensity PA, or both options combined. In addition, WHO recommends to "*Follow an online exercise class. Take advantage of the wealth of online exercise classes. Many of these are free and can be found on YouTube. If you have no experience performing these exercises, be cautious and aware of your own limitations*" (World Health Organization, 2020a). The use of exercise videos and eHealth, oriented to the encouragement and the delivery of PA through internet, social networks, TV, and mobile technologies are feasible options to maintain mental health and physical fitness during these important periods (Chen et al., 2020; Gunasekeran et al., 2021; Kraus et al., 2021; Tajudeen et al., 2021).

However, there is no data about the quality of the available eHealth and exercise videos, which is particularly relevant in this lockdown period due to COVID-19. Uploaded videos with health content coming from different sources present the risk of showing misleading and inaccurate information to users (Aydin & Akyol, 2020), and the authorship, the quality, and the validity of the information in the videos are essential topics to be considered (Atci, 2019; Brar et al., 2021, 2022; Güloğlu et al., 2022; M. Li et al., 2022; Osman et al., 2022; Şan, 2022).



0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, ...

(Fibonacci Sequence)

“Mathematics is the language of nature.”

Leonardo Fibonacci

3

3. Materials and Methods

3.1. Search Strategy

Several searches were conducted on *http://www.youtube.com/* using the different search terms which are also shown in such Table 1. The first 150 videos available to viewers were selected in the three articles included in this work. It was aimed to replicate a simple search strategy that could be conducted by any person or YouTube user. No restriction was applied to the search using filters; hence, YouTube sorted video results by their relevance according to the patented ranking algorithm active on each specific day. All the videos were added to a spreadsheet and then submitted for screening for duplicates, as well as to apply the exclusion criteria, showed in Table 1, by the research team.

Finally, after this process, 51, 133 and 68 videos were respectively included in each research (Article 1 -A1-, Article 2 -A2- and Article 3 -A3-) and independently viewed, analysed and evaluated by two different researchers in each case over a period of 5 weeks. These examiners were members of a research group at their corresponding universities with long and intensive experience in health topics research.

The search strategies are presented in Fig 1. which represents the flow of information through the different phases of a systematic review according to PRISMA (Moher et al., 2009).

#	Article 1	Article 2	Article 3
Title	"Quality Analysis of YouTube Videos Presenting Shoulder Exercises after Breast Cancer Surgery"	"Quality Analysis of YouTube Videos Presenting Pelvic Floor Exercises after Prostatectomy Surgery"	"Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown"
Search Date	31-Mar-19	24-Jan-21	03-Nov-20
Search Terms	Exercises after breast cancer surgery	Prostate cancer exercises pelvic floor	COVID Exercises Home
Exclusion Criteria	- Non-English language - Less than 5,000 views - Duplicated videos - Related with advertisements	- Non-English language - Duplicated videos - Pelvic floor exercises for women - Related with advertisements	- Non-English language - Videos that didn't show exercises - Duplicated videos - Related with advertisements
Videos included in quantitative and qualitative synthesis (meta-analysis)	51	133	68

Table 1. Search strategies for the articles included in the work.

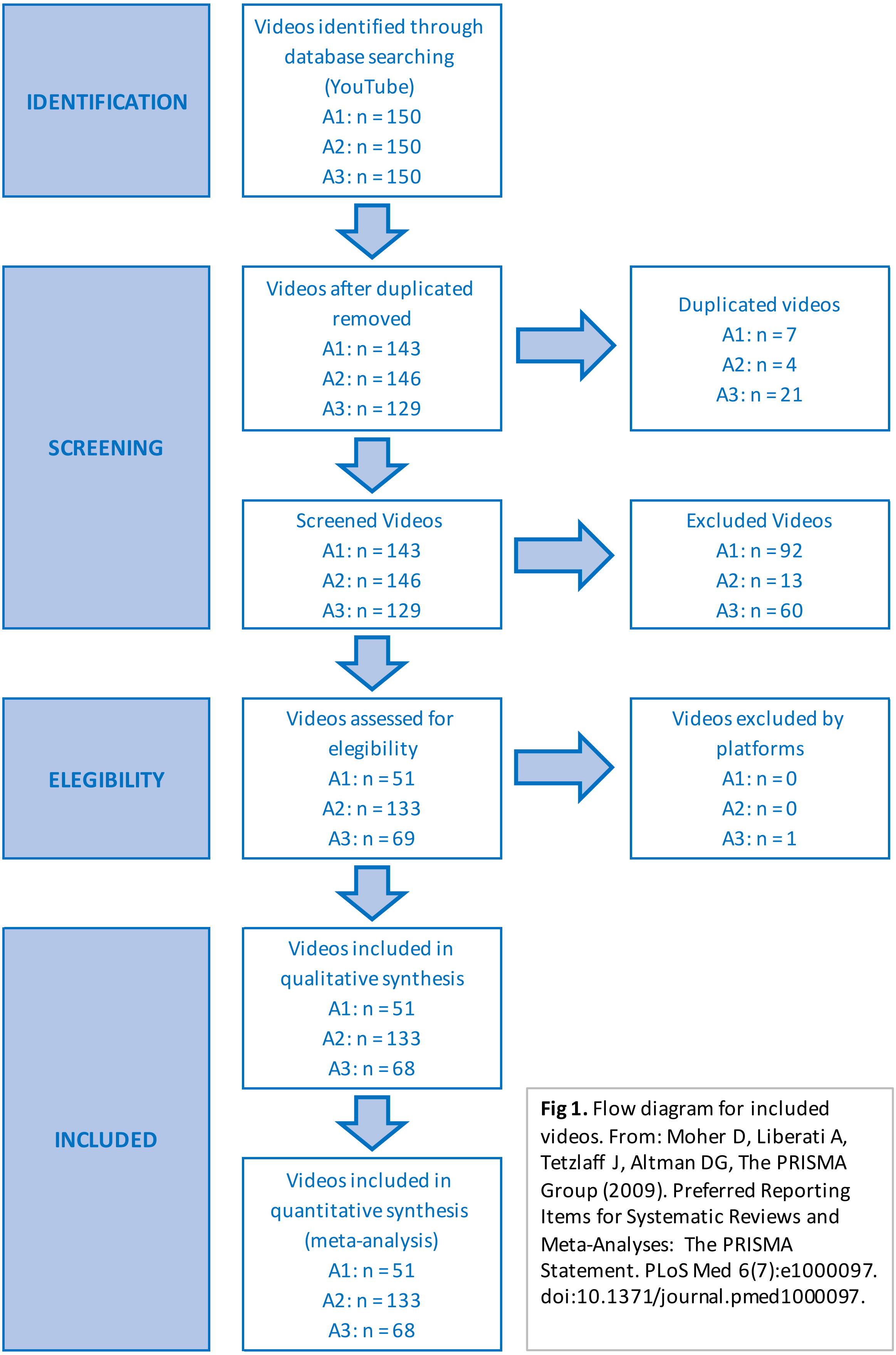


Fig 1. Flow diagram for included videos. From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7):e1000097. doi:10.1371/journal.pmed1000097.

Figure 1. Flow diagram.

3.2. Outcome Measures

Descriptive statistics were obtained from each video of the three datasets, calculating their mean and establishing their minimum and maximum values along with their standard deviation (SD). The variables that were included for this analysis were video length, number of views, days online, views ratio (views/day), number of Likes, number of Dislikes, number of Subscribers, Likes ratio (Likes/day), Dislikes ratio (Dislikes/day) and Video Power Index (in %). The same indicators were obtained for the quality scales mentioned before, which are Discern Score, Hon Score and GQS Score.

For the assessment of the concordance between examiners, the ICC was performed in the three articles. It is a widely used test–retest, intra-rater, and inter-rater reliability index, and it was conducted to gauge inter-examiner concordance with 95% CI based on mean rating (k = 2), consistency, two-way random model, and Pearson's correlation method. The level of significance was set at $p<0.05$.

According to different characteristics of the videos, they were categorized in different groups attending to their production source and continent of origin. Also, their descriptive characteristics were collected for each video: number of views (Views Count), Likes and Dislikes Count, number of days online and their publication date, and its duration (length in minutes and seconds).

As well, the number of exercises was collected in comparison with the corresponding *gold-standard* considered for each case:

3.2.1. Article 1

The scoring criteria for exercises were set according to recommendations by the Oncology Section of the American Physical Therapy Association (APTA) in collaboration with the American Cancer Society (ACS)(Cancer Research UK, 2020). Since seven exercises are recommended, the scoring system was based on a scale of 0 to 7 (1 point for each exercise coinciding with the 7 recommended by APTA) (American Cancer Society, 2019, 2021; Worthen A., 2022).

3.2.2. Article 2

Regarding to Prostate Cancer (PC), where the treatment involving the removal of the prostate may result in temporary or permanent erectile dysfunction (ED) and urinary incontinence (UI), pelvic floor exercises have been used to improve urinary continence following radical prostatectomy, with good results (de Lira et al., 2019; Hodges et al., 2020; Milios et al., 2019; Prota et al., 2012). Pelvic floor muscle training (PFMT) is one of the recommended techniques for the prevention, treatment, and rehabilitation of RP-related complications (de Lira et al., 2019). It can improve UI and ED after prostatectomy (Pan et al., 2019). The aim of PFME, first defined by Arnold Kegel in 1948 as a behavioural therapeutic method for treating incontinence (Kegel, 1948), is to enhance muscle volume and contraction strength in case of increased intra-abdominal pressure (Aydın Sayılan & Özbaş, 2018). According to these types of recommended exercises, different characteristics and indicators of the videos themselves were collected in the videos analysis in order to get much more information and being able to analyse which kind of exercises are being described and recommended in the videos.

3.2.3. Article 3

The World Health Organization (WHO) recommends, in a guide published on its website, for confined people without any respiratory illness or suspect of it, 150 min a week of moderated-intensity PA or 75 min a week of higher-intensity PA, or both options combined (World Health Organization, 2020a). Exercise time was calculated by multiplying days per week and the recommended exercise time on each video and was classified according to meeting WHO recommendations for exercises during lockdown (150 min/week of moderate exercise, 75 min/week of vigorous exercise or both combined). The number of exercises contained in each video that agreed with the 12 exercises during lockdown recommended by WHO were also collected (World Health Organization, 2020a).

For all the videos analysis carried out in the three articles, video popularity was assessed using the Video Power Index (VPI) (Atci, 2019; Aydin & Akyol, 2020; Gad et al., 2022; Gun et al., 2022), which can be defined according to the following formula:

$$VPI\ (\%) = \frac{Likes\ Count}{(Likes\ Count + Dislikes\ Count)} \times 100$$

Also, video popularity was assessed with the View Ratio (Braczynski et al., 2021; Cakmak & Mantoglu, 2021; McMahon et al., 2022; Pamukcu & Izci Duran, 2021) which is usually defined as follows:

$$View\ Ratio\ (n) = \frac{Views\ Count}{Days\ Online}$$

The educational quality of the selected videos was determined using the DISCERN instrument (Quality Criteria for Consumer Health Information) (Bahar-Ozdemir et al., 2021; Cakmak & Mantoglu, 2021; Charnock, 1998; Di Bello et al., 2022) and the Global Quality Scale (GQS) (Kunze et al., 2019; Marwah et al., 2021; Steeb et al., 2022).

The DISCERN questionnaire was created by the Division of Public Health and Primary Care at Oxford University, London, to gauge the quality of information regarding treatment choices for health problems and was first published in 1999 (Charnock et al., 1999). In the case of Article 1, the questionnaire consists of a total of 15 questions, in addition to an overall quality rating. Each question represents a different quality criterion and is rated from 1 to 5

(1: Very Poor, 2: Poor, 3: Average, 4: High, and 5: Very High quality). These 16 questions are classified into three sections: reliability (questions 1–8), information quality on treatment choices (questions 9–15), and overall score (question 16). Thus, the maximum total score is 80 points (Kılınç & Sayar, 2019). A modified 5-point DISCERN tool (Esen et al., 2019), adapted from the original DISCERN tool for assessment of written health information by Charnock et al., was used for Article 2 and Article 3 of this research. In these two cases, the questionnaire consists of a total of 5 questions in addition to an overall quality rating. Each question represents a different quality criterion, rated from 1 to 5 points (1: Very Poor, 2: Poor, 3: Average, 4: High, and 5: Very High quality), so the maximum total score is 25 points. In both versions of the questionnaire, higher scores indicate higher quality (Azak et al., 2022; Bahar-Ozdemir et al., 2021; Charnock et al., 1999; Memioglu & Ozyasar, 2022; Oydanich et al., 2022; Shungu et al., 2021)

GQS evaluates the content overall quality of online resources. One point is assigned for each of the 5 identifiable criteria present in a video, with 5 being the highest educational quality (Kunze et al., 2019). This scale covers accessibility and quality of the information, the overall flow of information, and how useful it would be for a user (Azak et al., 2022; Bahar-Ozdemir et al., 2021; Duran & Kizilkan, 2021; Erdem & Karaca, 2018; Esen et al., 2019; Memioglu & Ozyasar, 2022; Pamukcu & Izci Duran, 2021; Steeb et al., 2022).

All the videos in Article 1 and Article 3 were also evaluated with the HONCode tool, which was developed by the Health on the Net Foundation (Boyer, 2013; *HONcode: E-Guide for Health Consumers*, 2020) a non-profit organization accredited by the United Nations, who elaborated the code of conduct in order to standardize the reliability of online medical/health information (Boyer, 2013; Boyer et al., 1998). HONCode is the most frequently used assessment tool (Fahy et al., 2014) for reliability and credibility of the information that can be found on health-related websites and its certification is submitted to an annual review process made by the HON-Foundation, who also responds to any violation reported by internet users. The HONCode is not aimed to grade the information quality that is contained in a website, but rather establishes a series of rules in order for website publishers to comply with basic ethical standards of information delivery and to help ensure that visitors are always aware of the purpose and the source of the data they are viewing. It is constituted by a group of parameters about the reliability and credibility of the information that can be found in health-related websites. It is based on 8 principles determining the reliability of web pages and it can score from 0% to 100%. These principles are "*Authority*" and it checks the authors' qualifications; principle 2 checks the "*Complementarity*" regarding the information to support and not to replace; principle 3 is about "*Confidentiality*", regarding the respect the site users' privacy; principle 4 checks the "*Attribution*", i.e., the citation of the dates and sources of the medical information; principle 5 is about "*Justifiability*", that is, the justification of claims and if they are balanced and objective; principle 6 checks the "*Transparency*" regarding accessibility and the delivery of valid contact details; principle 7 is about "*Financial disclosure*" and it checks if funding details are provided. Finally, principle 8 checks the "*Advertising*", i.e., if it distinguishes advertisements in a clear way from the editorial matter (Boyer, 2013; Efe et al., 2021; Wilkens et al., 2022).

3.3. Statistical Analysis

Artificial intelligence (AI) is the study of complex information processing problems that have their roots in some aspect of biological information processing. The goal of the subject is to identify useful information processing problems and give abstract account of how to solve them. Such an account is called a "method", and it corresponds to a theorem in mathematics (Marr, 1977). There is growing interest in the application of AI techniques in bioinformatics. In particular, there is an appreciation that many of the problems in bioinformatics need a new way of being addressed given either the intractability of current approaches or the lack of an informed and intelligent way to exploit biological data (Hamet & Tremblay, 2017; Narayanan et al., 2002).

Machine Learning (ML) is a field within the area of AI that provides systems with the ability to learn and improve automatically, based on experience. These systems transform data into information, and with this information they can make decisions. Once the data is available, the learning process can begin. This process, carried out by an algorithm, tries to analyse and explore the data in search of hidden patterns. The result of this learning is sometimes nothing more than a function that operates on the data to calculate a certain prediction. The application of AI in medicine has two main branches: virtual and physical. The virtual component is represented by ML, (also called Deep Learning -DL-) that is represented by mathematical algorithms that improve learning through experience. There are three types of machine learning algorithms: (i) unsupervised (ability to find patterns), (ii) supervised (classification and prediction algorithms based on previous examples), and (iii) reinforcement learning (use of sequences of rewards and punishments to form a strategy for operation in a specific problem space) (Hamet & Tremblay, 2017; Mak & Pichika, 2019; Meskó & Górög, 2020; Panch et al., 2018). In other words, unsupervised learning focuses on discovering underlying structure or relationships among variables in a dataset and, once this is done, features discovered by unsupervised learning can often be incorporated into supervised learning models (Johnson et al., 2018). Like supervised learning, unsupervised machine learning methods exist on a continuum with more traditional statistical methods such as principal components analysis, mixture modelling, and various methods of clustering (Johnson et al., 2018). However, in recent years some new techniques that require fewer assumptions about the dataset have emerged, such as advanced algorithms for matrix or tensor factorization (Luo et al., 2017), topological data analysis (L. Li et al., 2015), and deep learning (Miotto et al., 2016).

Unsupervised learning fundamentally differs from supervised learning because it does not rely on specific classification instructions. Instead, it relies on autonomously grouping objects through exploration and discovery of the underlying pattern and structures in data. this approach is viewed as the primary way we as humans collect information and build knowledge about the world around us. Although this undirected approach may not afford unsupervised learning the ability to resolve the same level of detail for specific user-defined tasks, it provides a highly flexible approach to define patterns in data that do not need to be predicted or anticipated. Unsupervised learning also affords humans with the ability to handle unanticipated changes in conditions and extend knowledge outside the training parameters (Jha & Topol, 2016; Lecun et al., 2015; Roohi et al., 2020).

Factor Analysis (FA) was introduced in 1904 by Charles Edward Spearman (Spearman, 1904) and described in 1995 by Bartholomew D. J. (Bartholomew, 1995). This method makes it possible to create new variables from a set of original variables. It allows finding hidden (latent) causes that are a source of variability in the dataset (Jain & Dubes, 1988; Jolliffe, 2002). With this kind of latent variables, the number of variables can be reduced while keeping the maximum amount of information, and to set up a link between the observable causes and the new variables (known as factors or components).

Assuming that the input variables are correlated, the same amount of information can be represented with a smaller number of variables. In the derived solution, each original variable should be correlated with as few factors as possible, and the number of components should be kept to a minimum. The factor saturations reflect the influence of the k^{th} common factor on the j^{th} random variable (Bartholomew, 1995; Bartholomew et al., 2009; Ďuriš et al., 2021). There are several methods that can be used to estimate factor saturations, known as factor extraction methods.

The choice of an algorithm depends on a variety of factors that include, but are not limited to, data type/learning approach (supervised or unsupervised learning), the need for k (clustering), the importance of accuracy in the chosen model, the need for speed in data analysis, the data analysed, the size of the data set, the need for hierarchical output, and the need for categorical variables. Unsupervised ML methods are also sometimes referred to as clustering algorithms. Some of the most common methods employed in this approach include k-means clustering, anomaly detection, or certain statistical methods such as principal component analysis. These approaches usually utilize discrete or continuous data as their input parameter for identifying input regularities (i.e., k-means clustering) and/or for lowering dimensional representations (i.e., Principal Component Analysis) (Celebi et al., 2013; Chu et al., 2020; Rashidi et al., 2019; Yang et al., 2022).

As long as AI and ML techniques are broadly used in different fields of health care, such as policies definition (Bauer & Lizotte, 2021), , primary care research (Wang et al., 2021), prediction models in public health (Payedimarri et al., 2021), disasters and public health emergencies (Lu et al., 2022), Social Media and mental health (Laacke et al., 2021), medical imaging (Lei et al., 2021), cardiovascular health (Soto et al., 2021), etc., in this work, the Principal Components method is used for clustering applications. Other well-known methods are the Maximum Likelihood Method or the Least Squares Method. The number of new variables can be determined by the eigenvalue criterion (known as Kaiser´s rule), when factors that have their eigenvalues $\lambda>1$ are considered significant. Alternatively, it depends on the Scree Plot chart of eigenvalues, where the variance of each component is represented in order to the rest of them. The basic steps of FA are the selection of input data (assumption of correlation), the determination of the common factors, the estimation of parameters, the rotation of factors (Varimax Method—orthogonal rotation) and the factor matrix (factor saturation matrix) (Bro & Smilde, 2014; Ďuriš et al., 2021; Shahid et al., 2015; Yeung & Ruzzo, 2001)

Karl Pearson introduced the Principal Component Analysis (PCA) in 1901 for the first time (Pearson, 1901). The method pretends to transform the multidimensional input data in such a way that the output data are obtained from the most important linear directions, ignoring the less significant directions. Thus, the characteristic directions (characters) are extracted from the original data and, at the same time, the dimensions of the data are reduced.

The method is one of the basic methods of data compression: the original n variables can be represented by a smaller number of m variables, but explaining a significant part of the variability of the original data set. The collection of new variables (known as Principal Components -PCs-) consists of a linear combination of the original variables. The first principal component describes most of the variability of the original data set. The other PCs contribute to the overall variance, always with a smaller proportion. All pairs of PCs are perpendicular to each other (Bro & Smilde, 2014; Ďuriš et al., 2021; Gewers et al., 2021; Jolliffe, 2002). PCA and k-means are unsupervised methods that utilize discrete or continuous data as their input parameters for identifying input regularities (ie, clusters), so, their aim are clustering a medical data set to characterize subpopulations of a disease based on various known input clinical parameters and/or the dimensionality reduction of the dataset (Rashidi et al., 2019; Ryu et al., 2021).

The basic steps of the PCA include the construction of the correlation matrix from the original data, the calculation of the corresponding eigenvalues, the alignment from the major ($\lambda 1 > \lambda 2 > \cdots > \lambda n$), the calculation of the eigenvectors of the correlation matrix corresponding to its eigenvalues ($\upsilon 1, \upsilon 2, \ldots, \upsilon n$), the calculation of the variability of the original dataset ($\sigma 2$), the determination of the number of PCs that could represent the original variables based on variability and the transfer of the original data to a new reference base (Bro & Smilde, 2014; Ďuriš et al., 2021; Gewers et al., 2021).

According to Kaiser's Rule, using those PCs whose eigenvalue is greater than the mean of all eigenvalues (with standard data, the mean is 1, i.e. taking the PCs, whose eigenvalue is greater than 1) (Jolliffe, 2006b), we use the PCs which together explain the largest possible percentage of the total variance, or based on a graphical representation, the so-called Screen Plot graph, where it can be seen an inflection point and the PCs prior to this inflection point should be used (Ďuriš et al., 2021; Gewers et al., 2021; Shahid et al., 2015; Yeung & Ruzzo, 2001).

For the dataset of each of the articles of this work, three different statistical analyses were carried out: The first of them, which was modeled on techniques of machine automated learning, considers a problem of binary classification. It divides, using Principal Component Analysis (PCA) (Jain & Dubes, 1988; Z. Zhang & Castelló, 2017), the dataset into two groups: “*C1*” or “*Relevant Videos*” and “*C2*” or “*Not Relevant Videos*”. With PCA, the main information in a set of data can be visualized and described by multiple and interrelated variables. The information of any dataset is matched with the total variation. So, PCA finds some directions (the PCs) where the variance is maximized in the data and the clustering of the dataset points is done following the unsupervised ML algorithms (Fisher et al., 2022; Lau et al., 2022; Rodríguez et al., 2021; Ryu et al., 2021).

It is very useful for extracting the most important information from a set of data and it expresses the information referred to the new group of variables (or PCs). With this, PCA shrinks the number of dimensions of a set of data to a smaller number of dimensions or principal components.

For this study, two dimensions (PCA1 and PCA2) were enough to bundle the data of each article to be represented graphically, with a minimal forfeiture of the information of the original dataset (Beam & Kohane, 2018).

The aim of the first statistical test is to establish which variables distinguish in the best way the groups into which videos are divided in each article. The matter of this classification was checked in three situations according to these variables:

I. DISCERN: Videos classified in the "*C1/Relevant Videos*" group are the samples whose scores were bigger than the mean; the opposite was defining the second group, "*C2/Not relevant videos*".

II. Exercises number: Comparable with the preceding variable.

III. PCA: The binary variable defines both groups after a principal component analysis with a k-means algorithm. With this, *C1* and *C2* groups are defined by information variability with no loss of previous information. Accordingly, no information is lost when the dimensions are reduced to only two PCA, in these cases.

For these analyses, a grouping algorithm was used to obtain the least number of dimensions that best divides the videos into two different clusters. Precision was established by "*Leave-One-Out Cross Validation*" (LOOCV) (Vehtari et al., 2017) with the use of the "*Nearest Neighbor Classifier*" (NNC) (Kataria & Singh, 2008). Fisher's ratio (FR) was used to determine the variables' power of discrimination. The higher the FR is, the more discriminatory they are, which is a consequence of their low intraclass dispersion and their high interclass distance. In the case of a binary classification, the FR of a variable "j" is determined by:

$$\frac{(\mu_{j1} - \mu_{j2})^2}{{\sigma_{j1}}^2 + {\sigma_{j2}}^2}$$

where μ_{ji} is the measurement of the mass center of the distribution of probability of the variable "j" in group "i" (i = 1, 2), while σ_{ji} is the measurement of the dispersion inside that group.

Apart from FR, the Spearman, Kendall, and Pearson correlation factors were also calculated for them according to the determined classes. These additional factors disclose the importance, or discrimination power, of these variables according to the classification criteria (Bonett & Wright, 2000; Croux et al., 2010; Puth et al., 2015).

The second statistical analysis used the Wilcoxon test and the t-Test, showing the significance of the differences among both groups according to each variable´s point of view (Divine et al., 2013). This second test determines how relevant the variables are according to the classification of the videos within group "*C1*" and group "*C2*", for the three situations in each article. A t-Test establishes whether the mean of a variable in a group has a significant difference with respect to the other. The H_0, or null hypothesis, considers no such difference among these groups. If the result of the test is 1, it shows a big enough evidence that it rejects the Ho hypothesis, while if the result is 0, such H_0 will be accepted.

The accepted level of error probability or statistical significance is alpha $\alpha < 0.05$. Regarding the Wilcoxon test, it determines the difference among every set of couples and it tests this difference. According to this, its H_0, or null hypothesis, considers the equivalence between population medians for both videos´ groups "*C1*" and "*C2*".

In order to describe a given health state, many variables must be collected and analyzed, one by one, in numerical form among others, and look for correlations between them. This synthesis, although important in the descriptive analysis, does not take into account all the variables together. It remains fragmentary and gives an incomplete picture of the health status of interest. In order to remedy this shortcoming, the use of multivariate descriptive statistics allows a more accurate description of this state. PCA, which is a variant of factorial analysis, responds to this shortcoming and makes it possible to synthesize several variables together in order to describe as well as possible the set of individuals defined by these variables that are the subject of the descriptive analysis (Ben Salem & Ben Abdelaziz, 2021).

According to PCA, Henri Kaiser (1970) presented a Measure of Sampling Adequacy (MSA) for factor analytic data matrices that was subsequently adapted by Kaiser and Rice (1974) It is the function of the square of the elements of the matrix when they are compared to the original correlations´ squares (Kaiser, 1970, 1982). This factor, renowned as the Kaiser-Meyer-Olkin (KMO) index, is considered as "*Unacceptable*" if it is under 0.50, "*Miserable*" among 0.50 and 0.60, "*Mediocre*" if it is more than 0.60 and less than 0.70, "*Middling*" among 0.70 and 0.80, "*Meritorious*" for 0.80 to 0.90, and if it is more than 0.90 (and less than 1.00) it is classified as "*Marvelous*" (Kaiser, 1982). The KMO test was used to determine the multivariate normality and the sample adequacy. With the objective of evaluating the validity of the construct, the sample suitability for factor analysis was made using Bartlett's test of sphericity (Bartlett, 1951). Additionally, the statistical power of the carried-out analysis was checked through the combination of the PCA analysis and the KMO index (Jackson, 1991; Park, 2021).

The PCA outlook of the statistical assessment, this model can carry out statistical analysis on a vast number of observation data and reduce the dimension of this observation data without losing the main information of measured data. Starting from the correlation coefficient matrix of observation data, several comprehensive factors that can reflect the main information of the original data and control all data are obtained (Shihui Zhang et al., 2022).

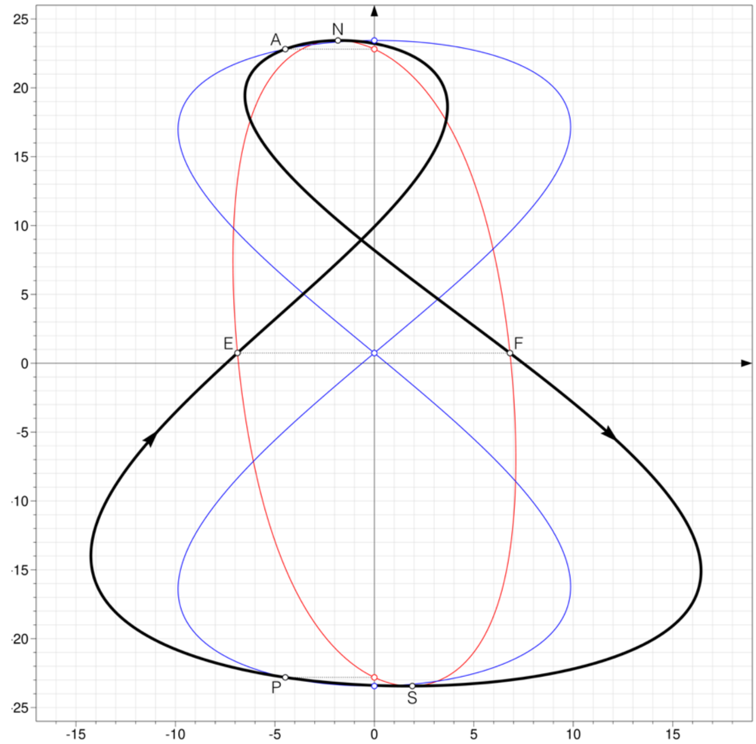
PCA is a statistical method that uses the eigenvector to determine the orientation of features. PCA's fundamental idea is to map a j-dimensional feature space into an i-dimensional space, which is generally known as the PCs, where i < j (Dweekat & Lam, 2022). For each dataset, the covariance matrix was calculated, and the eigenvectors and eigenvalues were computed. Because an eigenvalue shows the most significant relationship between the dataset characteristics, the eigenvector with the greatest eigenvalue is selected as the principal component of the datasets. The eigenvalues were sorted in ascending order to select the most significant principal component(s), while the lowest eigenvalues were discarded. This process reduces large dimensional datasets to smaller dimensional datasets. The variance measures the dispersion of the data in the dataset. Lastly, eigenvectors and eigenvalues for the covariance matrix were computed. Eigenvalues were transformed using the varimax orthogonal rotation method (Salih Hasan & Abdulazeez, 2021). PCA can also remove redundant information in the fused features of big datasets (Yu et al., 2022).

The statistical power of these two combined analyses indicates that the PCA analysis allows the videos to classify themselves into two groups (C1, high-quality videos and C2, low-quality videos) without any information being lost on any of the variables considered in the entire study. Thus, the principal components of the PCAs show a high statistical relationship with the variables that evaluate the quality of the videos, which enables the quality of the videos to be analyzed through the two principal components, instead of it having to be done through all the variables of the study (Kaiser, 1970, 1982; Salih Hasan & Abdulazeez, 2021; Z. Zhang & Castelló, 2017) This is a huge advantage when carrying out this kind of analysis, considering that any important information is not lost through these two principal components because their level of relationship with the main variables that evaluate the quality of the videos is maximum and p values reveal a high, statistically significant relation (Kambhatla & Leen, 1997; Salih Hasan & Abdulazeez, 2021; Yu et al., 2022; T. Zhang & Yang, 2016; X. Zhang et al., 2022).

The Kaiser–Meyer–Olkin measure for the adequacy of the sample (KMO index) for the PCA was performed to check the sample size compliance before applying the analysis of factors (Pechenizkiy et al., 2004). The KMO measure is a test that is used to decide whether a group of samples are suitable for conducting factor analysis, and it is calculated in terms of the correlation and partial correlation between the variables. The result of this test showed that the KMO value reached 0.916 in Article 1, 0.872 in Article 2 and 0.845 in Article 3 (Table 11). According to this, it can be determined that the sizes of the samples were, according to Kaiser and its KMO indicator, “Meritorious” for these datasets. This KMO measure also indicates the power of statistical analysis that was carried out. According to this, the statistical power of these combined analyses is such that it allows the videos classification, according to their quality, by considering two variables (principal components) that combine together the information of all the other variables analyzed in the study with no loss of information and removing redundant information (Gewers et al., 2021; Mahmoudi et al., 2021; Saerens et al., 2004; Yu et al., 2022).

So, according to this, as PCA is the most used dimensionality reduction technique (Kambhatla & Leen, 1997; Zebari et al., 2020; T. Zhang & Yang, 2016) if it is considered a graphical representation of the coordinates of each video in a coordinate system where the axes represent the principal components (i dimensions) (Jolliffe, 2006a), it can be seen in the following graphs that videos are grouped naturally and clearly, with no bias or manual intervention, in two clusters of points. These two clusters include those videos with higher scores in the quality variables, such as GQS, DISCERN and HONCode, in one cluster (C1), and videos with lower scores in the other cluster (C2).

For the assessment of the concordance between examiners, the intraclass correlation coefficient (ICC) analysis was performed with confidence intervals (CI) of 95% considering a two-way random model, mean rating (k=2) consistency, and the method of Pearson's correlation. The significance level was fixed at $p < 0.05$. With a confidence interval of 95%, values resulting from the ICC calculation under 0.5 mean a "*Poor*" reliability, "Moderate" for values between 0.5 and 0.75, "*Good*" between 0.75 and 0.90, and "*Excellent*" reliability for values higher than 0.90 (Bartlett, 1951; Koo & Li, 2016). Also, the coefficients of Pearson's correlation for the main variables referred to DISCERN score were also calculated. These calculations show the discrimination power (relevance) of the rest of the indicators considering a level of significance of $p<0.05$.



Analemma of the Sun

(An analemma is a diagram showing the position of the Sun in the sky as seen from a fixed location on Earth at the same mean solar time)

"Keep your eyes on the stars,

and your feet on the ground"

Theodore Roosevelt

4

4. Results

4.1. Inter-reviewers' concordances

For the assessment of the concordance between examiners, the ICC was performed in the three articles. With the mentioned confidence interval of 95%, and the level of significance set at $p<0.05$, values resulting from the ICC calculation under 0.5 mean a "Poor" reliability, "Moderate" for values between 0.5 and 0.75, "Good" between 0.75 and 0.90, and "Excellent" reliability for values higher than 0.90 (Koo & Li, 2016), as it is detailed in Table 2.

Inter-reviewer agreement, for the Articles included in this study, scored 0.9860 in Article 1, 0.9505 for Article 2 and 0.9299 for Article 3.

INTERVAL	RELIABILITY *
Less than 0.5	Poor
0.5 - 0.75	Moderate
0.75 - 0.90	Good
Greater than 0.90	Excellent
** Based on the 95% confidence interval*	

Table 2. ICC Intervals

4.2. Descriptive Statistics

After collecting the data from the videos, the lists were assigned to two researchers, in each case of the three articles, for the analysis of the videos. ICC analysis was performed for each articles´ results and, subsequently, the average of both researchers' evaluations was used for the analysis of the common information and results, i.e., the data used for each video were the mean of the scores assigned by each of the independent researchers, including the HONCode, GQS, and DISCERN scores.

As explained before, in Article 1, DISCERN scores over a maximum of 80 points while in Article 2 and Article 3, it scores over a maximum of 5. GQS scores over a maximum of 5 in all of the three articles. HONCode scores over a maximum of 100 also in all of the three articles that compose this research:

- According to these scores, the DISCERN mean was:
  - 50.97 (SD 10.19) for Article 1,
  - 3.35 (SD 1.25) for Article 2
  - and 2.29 (SD0.71) for Article 3.

- HONCode scored a mean of:
  - 78.30 (SD 11.02) for Article 1,
  - HONCode was not available for the research in Article 2
  - and 58.95 (SD 12.89) for Article 3.

- GQS had a mean of:
  - 3.49 (SD 0.74) for Article 1,
  - 3.38 (SD 1.02) for Article 2
  - and 2.32 (SD 0.86) for Article 3.

Descriptive statistic basic data of the mean results for VPI, HONCode, GQS, and DISCERN are shown in Table 3.

Descriptive Statistics	Mean ± Standard Deviation		
	Article 1	Article 2	Article 3
Video length (min:sec)	9:42 ± 9:15	14:42 ± 19:34	10:37 ± 10:55
Views (n)	53,963 ± 67,376	124,354 ± 472,419	59,565 ± 229,286
Days online (n)	2,158 ± 922	1,777 ± 1,180	42.3 ± 13.6
View ratio (views/days)	27.14 ± 30.24	138.30 ± 788.31	1,326.62 ± 4,860.72
Likes (n)	245.0 ± 320.5	1,082.0 ± 4,883.0	2,275.0 ± 9,575.0
Dislikes (n)	13.4 ± 14.2	68.6 ± 265.5	29.0 ± 105.0
Subscribers	226,793 ± 197,973	95,039 ± 343,699	238,278 ± 680,369
Likes/day (n)	0.15 ± 0.22	1.65 ± 11.35	50.75 ± 205.78
Dislikes/day (n)	0.008 ± 0.013	0.090 ± 0.520	0.670 ± 2.290
VPI (%)	93.48 ± 5.42	92.28 ± 8.89	97.33 ± 3.66
DISCERN score	50.97 ± 10.19	3.35 ± 1.25	2.29 ± 0.71
HON score	78.30 ± 11.02	-	58.95 ± 12.89
GQS score	3.49 ± 0.74	3.38 ± 1.02	2.32 ± 0.86

Table 3. Descriptive Statistics of the videos included in the studies.

Table 4 shows the mean results for the DISCERN Scores according to the source of production (author) and continent of origin in each of the articles of this research.

Mean DISCERN Scores		Article 1	Article 2	Article 3
SOURCE	Academic Institutions	56.25	3.88	3.25
	Health Institutions	54.85	3.88	2.83
	NGOs	51.38	4.33	2.45
	Media	50.94	2.79	2.50
	Individuals	44.78	1.75	2.68
	TOTAL	**50.97**	**3.35**	**2.29**
ORIGIN	America	51.73	3.33	2.20
	Europe	51.57	3.40	2.58
	Asia	47.00	3.00	2.25
	Australia	35.25	3.55	3.50
	Africa	0.00	0.00	2.00
	TOTAL	**50.97**	**3.35**	**2.29**

Table 4. Mean DISCERN Scores

As mentioned before, while Article 1 used the extended version of the DISCERN Scale, Article 2 and Article 3 used the reduced version of this scale. So, with the purpose of easing the comparison of these results, a scale-change was applied to the results of Article 1 from a 0-to-80 scale to a 0-to-5 scale. The results are shown in Table 5 where they can be compared in an easier way. In this table, it can be seen that, in Article 1 and Article 3, Academic Institutions get the highest score (3.52 and 3.25, respectively) and Health Institutions get the second highest score (3.43 and 2.83, respectively). In Article 2, these two kinds of institutions get the same score (3.88, both of them) which is the second highest one, while NGOs get the highest score (4.33). Media and Individuals get the lowest scores in the three articles and, only in the third one, NGOs are included also in the lowest results group.

Regarding the origin of the videos, i.e. the continent where they were produced, in Article 1 the highest score is obtained by the videos produced in America (3.23), closely followed by those produced in Europe (3.22). In Article 2 and Article 3, videos produced in Australia get the highest score (3.55 and 3.50, respectively), followed by the videos produced in Europe (3.40 and 2.58, respectively). In the three articles, videos produced in Africa get the lowest scores of all (0.00, 0.00 and 2.00, respectively).

Mean DISCERN Scores		Article 1* (base changed)	Article 2	Article 3
SOURCE	Academic Institutions	3.52	3.88	3.25
	Health Institutions	3.43	3.88	2.83
	NGOs	3.21	4.33	2.45
	Media	3.18	2.79	2.50
	Individuals	2.80	1.75	2.68
	TOTAL	**3.19**	**3.35**	**2.29**
ORIGIN	America	3.23	3.33	2.20
	Europe	3.22	3.40	2.58
	Asia	2.94	3.00	2.25
	Australia	2.20	3.55	3.50
	Africa	0.00	0.00	2.00
	TOTAL	**3.19**	**3.35**	**2.29**

Table 5. Mean DISCERN Scores (after the base change in Article 1)

Table 6 shows the mean results for the GQS Scores according to the source of production (author) and continent of origin in each of the articles of this research. The same GQS scoring system was used for all the articles, so there is no necessity of any base change to compare the results. All of them are presented on the 0-to-5 scale defined by the original GQS methodology. In the table below, Table 6, it can be seen that Academic Institutions, Health Institutions and HGOs get the highest scores (in different order, but always these three types of authors) in the GQS scale. Academic Institutions get the highest score in Article 1 (3.83) and in Article 2 (4.00, same result as NGOs), followed in both cases by Health Institutions (3.74 in Article 1 and 3.77 in Article 2). In Article 3, Health Institutions get the highest score (3.33) followed by NGOs and Academic Institutions, which obtained close scores (3.06 and 3.00, respectively). In the three articles, the lowest scores were obtained by Individuals (3.16, 2.82 and 1.84, respectively) followed by Media (3.31, 2.96 and 2.42, respectively).

Regarding the origin of the videos, in Article 1 videos produced in Europe received the highest scores (3.79) followed by those produced in America (3.52). As in the DISCERN Scale happened, in Article 2 and Article 3, videos produced in Australia received the highest scores (3.55 and 3.50, respectively), followed by those produced in Europe (3.40 and 2.54, respectively). And, as it also happened in the DISCERN scores, videos produced in Africa received the lowest scores, which are the same that these videos obtained in the DISCERN analysis (0.00, 0.00 and 2.00, respectively).

Mean GQS Scores		Article 1	Article 2	Article 3
SOURCE	Academic Institutions	3.83	4.00	3.00
	Health Institutions	3.74	3.77	3.33
	NGOs	3.63	4.00	3.06
	Media	3.31	2.96	2.42
	Individuals	3.16	2.82	1.84
	TOTAL	**3.49**	**3.38**	**2.32**
ORIGIN	America	3.52	3.36	2.24
	Europe	3.79	3.40	2.54
	Asia	3.50	3.00	2.30
	Australia	1.75	3.55	3.50
	Africa	0.00	1.00	2.00
	TOTAL	**3.49**	**3.38**	**2.32**

Table 6. Mean GQS Scores

Considering the base changed for DISCERN Scores in Article 1, which converts them from the original scale from 0 to 80 to the modified scale, which goes from 0 to 5, the quality level of the videos can be ranked as shown in Table 7. The same quality intervals apply to the scores obtained by the videos in the GQS scale (Charnock et al., 1999; Erdem & Karaca, 2018; Esen et al., 2019; Kunze et al., 2019; Yalkin et al., 2022).

DISCERN and GQS Interval	Quality Level
0 - 1	Very Poor
1 - 2	Poor
2 - 3	Average
3 - 4	High
4 - 5	Very High

Table 7. DISCERN and GQS quality levels intervals

Regarding the HONCode results, in Article 1, Academic Institutions received the highest score (62.96), followed by NGOs (28.78), and Media the lowest one (10.40). In Article 3, Health Institutions received the highest score (72.33), followed by NGOs (67.43) and Academic Institutions (66.63). and Individuals the lowest punctuation (53.06).

According to the continent of origin, in Article 1 the videos with better punctuation are those from Asia (50.69) followed by videos from Europe (45.03). In Article 3, the videos from Australia received the highest score (69.00).

As explained before, the HONCode instrument was not available for Article 2, which is the reason why there are no results for such tool for this second Article. All these mentioned results are shown in Table 8.

Mean HONCode Scores		Article 1	Article 2	Article 3
SOURCE	Academic Institutions	62.96	-	66.63
	Health Institutions	24.95	-	72.33
	NGOs	28.78	-	67.43
	Media	10.40	-	58.50
	Individuals	26.24	-	53.06
	TOTAL	**27.14**	**-**	**58.95**
ORIGIN	America	23.75	-	56.80
	Europe	45.03	-	62.73
	Asia	50.69	-	61.85
	Australia	20.61	-	69.00
	Africa	0.00	-	63.00
	TOTAL	**27.14**	**-**	**58.95**

Table 8. Mean HONCode scores

In the following table, the Pearson´s correlation coefficient of GQS and HONCode regarding to DISCERN are shown (Table 9).

Pearson's correlation coefficient of GQS and HONCode regarding to DISCERN			
Article	**Variable**	**Pearson coefficient***	**p value****
Article 1	HONcode	0.700	< 0.001
	GQS	0.650	< 0.001
Article 2	HONcode	-	-
	GQS	0.902	< 0.001
Article 3	HONcode	0.606	< 0.001
	GQS	0.916	< 0.001

* Pearson's correlation coefficients (t-test—2 tailed)

** α = 0.05

Table 9. Pearson's correlation coefficients

Considering the number of exercises in each article regarding the DISCERN Scores it can be seen in Table 10 that, the higher result videos get in the DISCERN scale, the higher is the number of recommended exercises. The mean number of recommended exercises is increasing from the "Very Low" score in the DISCERN scale with 0.00, 0.61 and 1.05 respectively for Article 1, Article 2 and Article 3; to 1.20, 1.27 and 1.67 (respectively) for "Low" quality videos; to 1.45, 2.06 and 2.33 for "Average" quality videos; 2.10, 2.48 and 2.83 for "High" quality videos and 1.00 and 3.58 for "Very High" quality videos (no video scored "Very High" in the DISCERN scale in Article 3). Regarding the GQS Scores, also the number of recommended exercises is growing with higher scores, as can be seen in Table 10, unless in the already mentioned case of Article 3 where no video scored "Very High" neither in DISCERN nor GQS scales. Also, in Article 1 the number of recommended exercises for "Very High" score in DISCERN scale is 1.00 and it seems to be out of tendency as it also happens with "Average" score in the GQS scale (1.10 in Article 1). To soften these out-of-trend situations, if we consider the average score of the three articles on each scale, average of averages, we obtain a clear and clean increasing tendency of the number of recommended exercises as the score on the DISCERN and GQS scales increases (Table 10).

Mean Number of Exercises				
DISCERN Score	**Article 1**	**Article 2**	**Article 3**	**Mean**
Very Low	0.00	0.61	1.05	0.55
Low	1.20	1.27	1.67	1.38
Average	1.45	2.06	2.33	1.95
High	2.10	2.48	2.83	2.47
Very High	1.00	3.58	-	2.29
Total	**1.75**	**1.93**	**1.72**	**1.80**
GQS Score	**Article 1**	**Article 2**	**Article 3**	**Mean**
Very Low	0.67	0.73	1.00	0.80
Low	1.33	1.59	1.57	1.50
Average	1.10	2.13	2.80	2.01
High	2.30	2.74	2.67	2.57
Very High	7.00	3.08	-	5.04
Total	**1.75**	**1.93**	**1.72**	**1.80**

Table 10. Mean Number of Exercises regarding Discern and GQS Scores

The graphical representation of the PCs coordinates (Jolliffe, 2006a) of the videos, being each video represented by a point on a system of coordinate axes where each axis is one of the principal components, allows a clear visualization of the self-organization to which the videos are subjected without any manual intervention. The clear differentiation of the points clouds in which the high-quality videos are grouped on one hand and the low-quality videos on the other hand, clearly show that this naturally occurring grouping allows an intuitive visualization of the quality levels of each of the points clouds. Also in Table 11, with the values of the KMO indicators, the total explained variances of the PCA for each article are shown. The values of the explained variances indicate that most of the information of the original datasets is included in the dimensional reduction to the new reference system whose bases are the calculated PCs as a linear-combination of the original variables in each dataset, respectively.

Article	KMO	Total Explained Variance
Article 1	0.916	86.028%
Article 2	0.872	95.889%
Article 3	0.845	93.268%

Table 11. KMO values and Total Explained Variance for each article.

A Scree Plot displays the explained variation for each PCs in descending order versus the number of the components, which is generally used in PCA analysis to choose the number of PCs (Horn, 1965; Ivosev et al., 2008; Kanyongo, 2005). As mentioned before, only the PCs whose eigenvalue is higher than 1 are considered as valid PCs for being considered in the rest of the calculation process described by the PCA methodology (Abdi & Williams, 2010; Ďuriš et al., 2021; Eastment & Krzanowski, 1982; Gewers et al., 2021; Shahid et al., 2015; Yeung & Ruzzo, 2001). As it can be seen in the following figures (Figure 2, Figure 4 and Figure 6), the corresponding Scree Plots for each article shows the accumulated explained variance level for each of the PCs (Hasan & Tahir, 2010; Ledesma et al., 2015; Qu et al., 2002; Silva et al., 2020; Z. Zhang & Castelló, 2017; Zhu & Ghodsi, 2006).

4.3. Article 1

Figure 2 shows the Eigenvalues for Article 1 which denote variance accounted for by a linear combination. As can been seen in the Figure, the amount of variance accounted is greatest for the first component, followed by the second, but diminishes greatly thereafter; this reveals the 2-component solution is the best for this case because only the first two PCs present Eigenvalues greater than 1. With this PCs and the KMO indicator of 0.916, the accumulated explained variance for these two PCs is 86.028%.

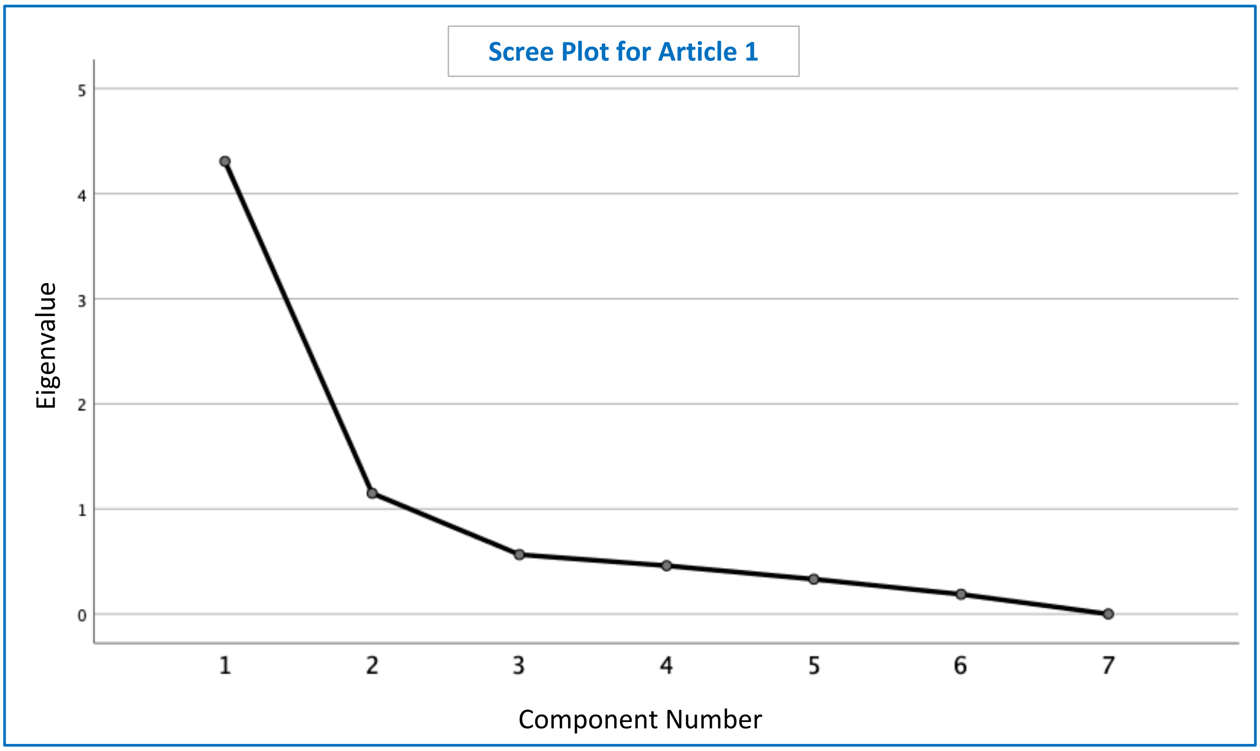


Figure 2. Scree Plot for Article 1

The PCA component diagram of Article 1, representing the naturally formed groups of videos with similar quality levels, is plotted in the figure below (Figure 3). High quality videos are represented in blue color and low quality videos are represented in pink color. As can be seen, there is a clear differentiation between the groups of videos of both qualities and these are naturally grouped in accordance to their coordinates respect to the reference frame of the PCs.

Principal Components Analysis (PCA)

(Varimax Rotation Method with Kaiser Normalization)

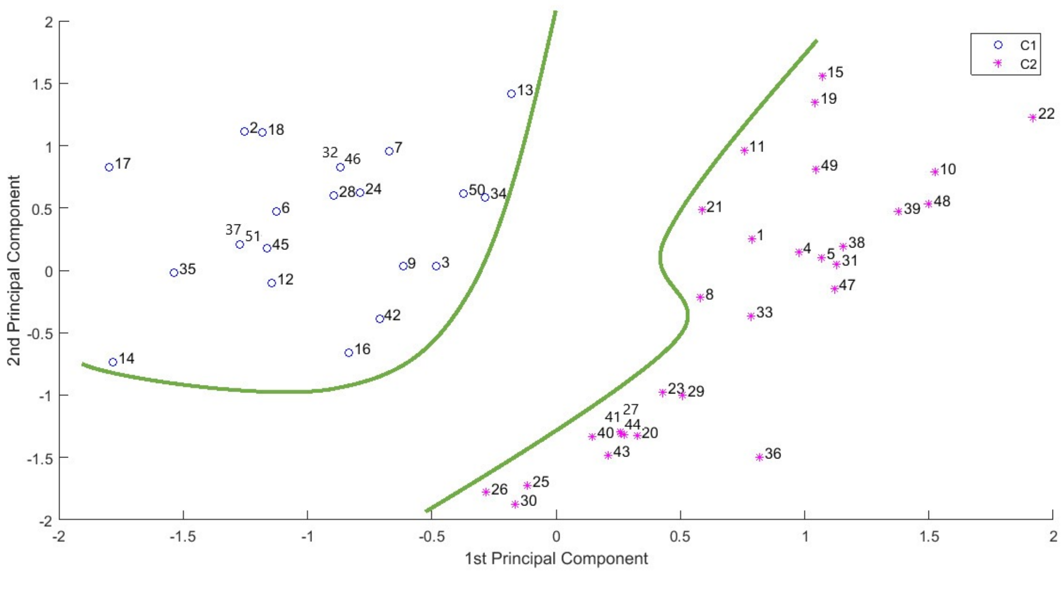


Figure 3. PCA Components Diagram for videos in Article 1.

4.4. Article 2

Figure 4 shows the Eigenvalues for Article 2 which denote variance accounted for by a linear combination. As can been seen in the Figure, the amount of variance accounted is greatest for the first component, followed by the second, but diminishes greatly thereafter; this reveals the 2-component solution is the best for this case because only the first two PCs present Eigenvalues greater than 1. With this PCs and the KMO indicator of 0.872, the accumulated explained variance for these two PCs is 95.889%.

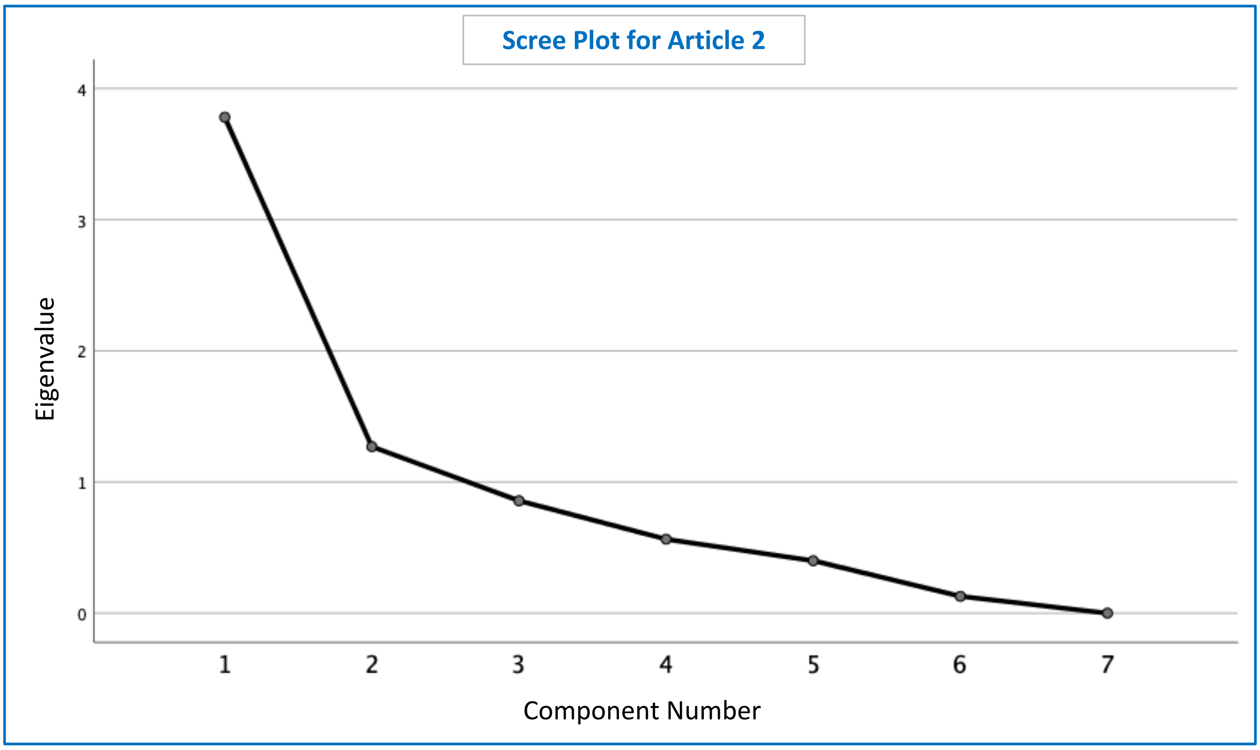


Figure 4. Scree Plot for Article 2

The PCA component diagram of Article 2, representing the naturally formed groups of videos with similar quality levels, is plotted in the figure below (Figure 5). High quality videos are represented in blue color and low quality videos are represented in pink color. As can be seen, and similar to the previous case, there is a clear differentiation between the groups of videos of both qualities and these are naturally grouped in accordance to their coordinates respect to the reference frame of the principal components.

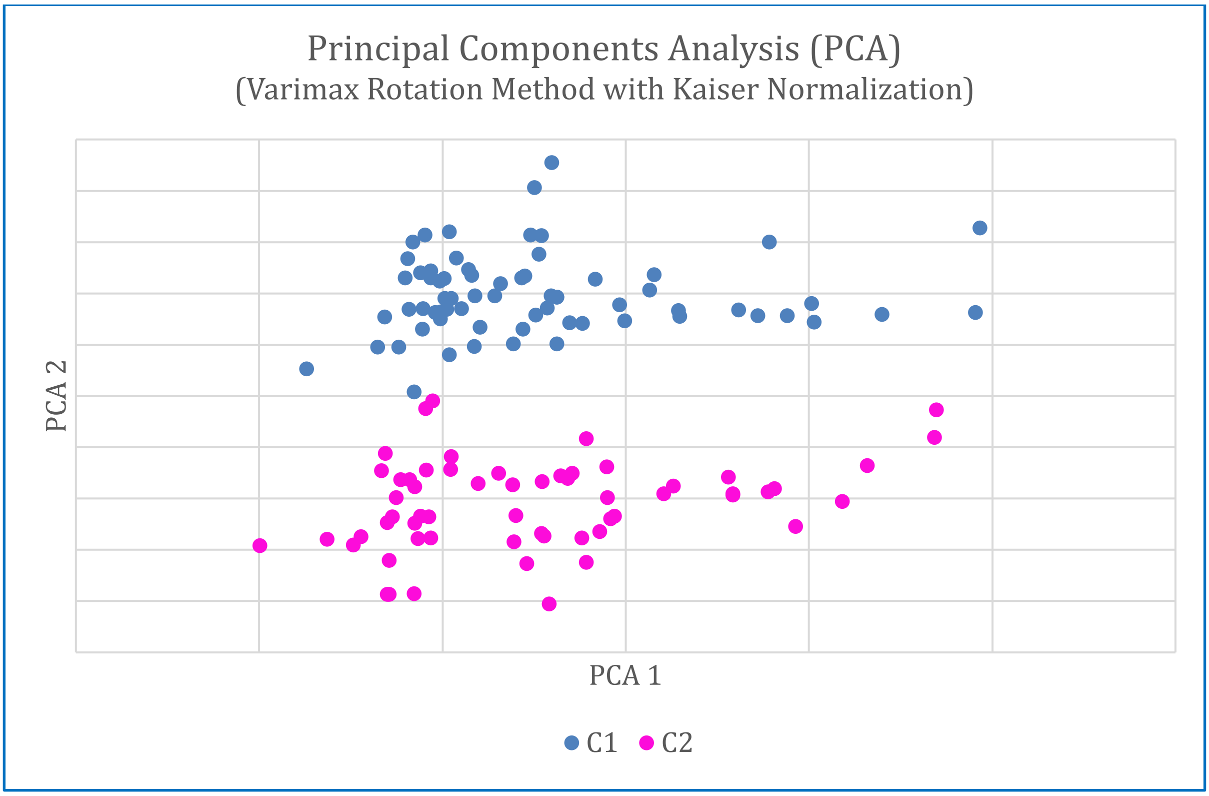


Figure 5. PCA Components Diagram for videos in Article 2.

4.5. Article 3

Figure 46shows the Eigenvalues for Article 3 which denote variance accounted for by a linear combination. As can been seen in the Figure, the amount of variance accounted is greatest for the first component, followed by the second, but diminishes greatly thereafter; this reveals the 2-component solution is the best for this case because only the first two PCs present Eigenvalues greater than 1. With this PCs and the KMO indicator of 0.845, the accumulated explained variance for these two PCs is 93.268%.

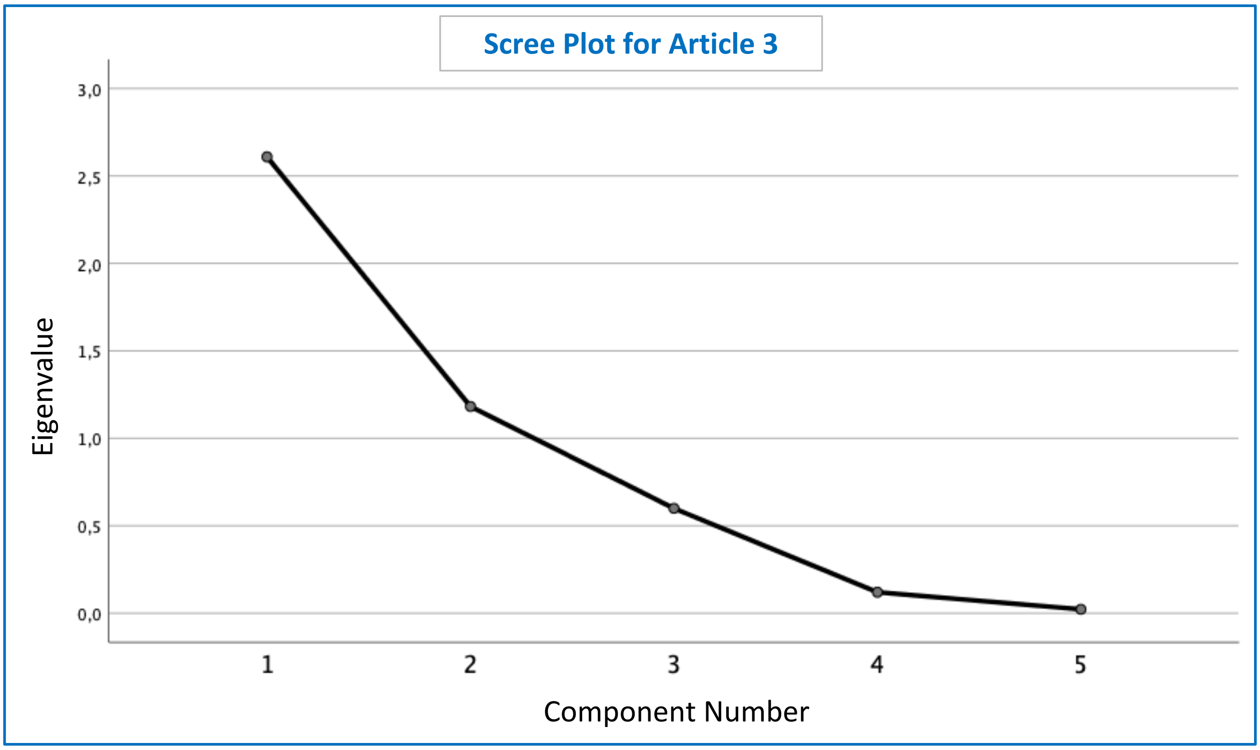


Figure 6. Scree Plot for Article 3

The PCA component diagram of Article 3, representing the naturally formed groups of videos with similar quality levels, is plotted in the figure below (Figure 7). As in both previous articles, high quality videos are represented in blue color and low quality videos are represented in pink color. As can be seen, and similar to both previous cases, there is a clear differentiation between the groups of videos of both qualities and these are naturally grouped in accordance to their coordinates respect to the reference frame of the principal components.

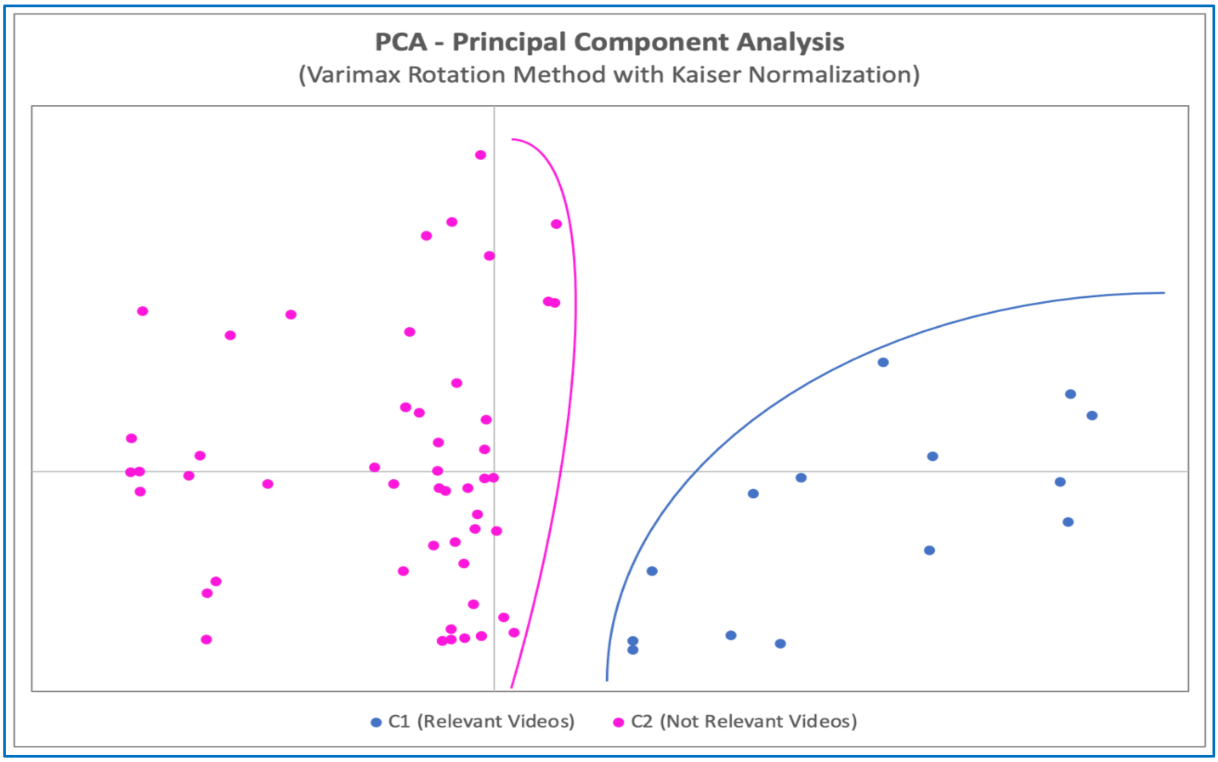
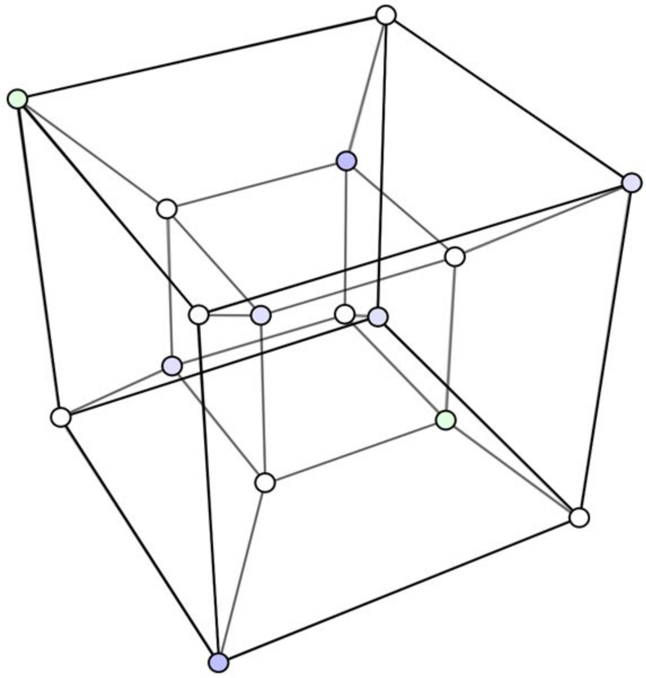


Figure 7. PCA Components Diagram for videos in Article 3.



Tesseract

(Hypercube of 4-dimmensions)

"A good discussion increases the dimensions

of everyone who takes part"

Randolph Bourne

5

5. Discussion

This doctoral thesis was conducted in order to better comprehend the nature of evidence and information independently accessed by patients and any other kind of user on a massive online, media-sharing platform which use is daily increasing. The growing use of new technologies and vast amount of content available have turned the internet into an important source of health information for healthcare professionals (who use new technologies more frequently to communicate with patients and achieve the necessary adherence and interaction in health processes) as well as for patients (who use it as a source of information). YouTube is one of the most important development tools of the eHealth era. Videos are a simple, attractive, and affordable source of information, with vast dissemination capacity available.

However, this entails some risk, in as much as the information available on the web is not subject to any kind of inspection. Consequently, this can lead to erroneous and even harmful messages being conveyed to users and/or viewers. In the search procedure to improve health processes, rigorous education and dissemination are necessary, preferably carried out by health professionals.

In the case of the analyzed videos, most of them (33.0% and 32.3% in Articles 1 and 2) were produced by Health Institutions, which can be considered as a guarantee, due to the fact that they are in the top places of the quality scales. In the case of Article 3, videos produced by Health Institutions represented a small percentage of the total but also reached very high scores in the quality analysis, which can be considered as a guarantee of the information shared in those videos.

ICC is a widely used test–retest, intra-rater, and inter-rater reliability index. Based on the 95% confidence interval of the ICC estimation (Koo & Li, 2016), the inter-reviewer agreement for these studies were, respectively for the three articles, 0.9860, 0.9505 and 0.9299, pointing towards an "*Excellent*" concordance between the two examiners that participated in each of the three articles that integrate this study, according to the ICC intervals shown in Table 2.

The average lengths of the videos in our studies were 9:42 for Article 1, 14:42 for Article 2 and 10:37 for Article 3, which is aligned with previous studies that reported an average duration of 6:17–10:35 min (Gokcen & Gumussuyu, 2019; Kuru & Erken, 2020), except in Article 2 which has the bigger standard deviation and also 63.91% of the videos have a duration of less than 10:00 min, which points a high similitude with previous studies.

No statistically significant relationship was found between video length and their quality or popularity, contrary to the study by Akif Aydin that found that videos with the longest duration and the highest VPI appeared to be associated with higher quality scores (Aydin & Akyol, 2020).

While the most frequently used approach for reduction of dimensionality is PCA, it, in essence, reduces the high number of dimensions in a large dataset to fewer dimensions to speed up the storage and processing procedure, which makes the data more interpretable and faster to be processed and analyzed. It is a statistical technique that preserves the largest amount of information and eliminates redundant noise and data.

After performing these two statistical analyses, the PCA calculation with the KMO index allows the videos to auto-classify into two clusters (C1 and C2, High- and Low-Quality videos, respectively) without losing any information of any variable of the study. All the variables were combined, through the PCA methodology, into two new variables (Principal Components) that contain all the information of the original variables. According to this, the quality of the videos can be assessed attending only to these new variables, which makes the assessment much quicker, easier, and with no loss of information, which represents an enormous advantage when performing this kind of analyses, also because the *p*-values show a relationship with high statistical significance.

The statistical strength of these two analyses is so great that it permits the self-classification of the videos, considering all their characteristics (quality among them), using only two variables (the principal components) which incorporate the data of the totality of the variables considered in this research. One of the most important and relevant considerations of this analysis is that no manual intervention in made in the dataset, and that is one of the reasons why this statistical technique is highly valuable when analysing large amounts of data, due there is no manual intervention and errors and bias are avoided.

Although the number of recommended exercises could be higher, the video scores are "*High*" and "*Average*" on the quality scales that were considered for these articles and this research, which is in line with other authors (Akyol & Karahan, 2020; Altan Şallı et al., 2020; Karakoyun & Yildirim, 2021; Meteran et al., 2022; Morra et al., 2022; Nangia et al., 2020; Navarro et al., 2021; Onder & Zengin, 2021; Salman & Bayar, 2021).

Regarding the number of exercises in each article, the results of these analyses are the following;

5.1. Article 1

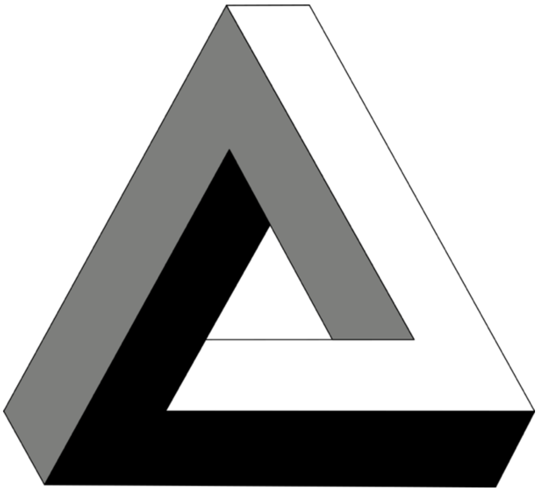
The Oncology Section of the American Physical Therapy Association (APTA), in collaboration with the American Cancer Society (ACS), recommends a series of seven exercises. Out of these seven, the analyzed videos show a mean number of 1.75 exercises (which represents the 25% of recommended exercises) and, videos that scored High quality in the quality scales show a 30% of the recommended exercises (mean of 2.10 exercises). There were some videos that showed from five to seven exercises, out of those seven recommended ones.

5.2. Article 2

The mean number of recommended exercises, according to standards, is 1.93 and Very High quality videos show a mean of 3.58 exercises.

5.3. Article 3

Out of the 12 suggested exercises by WHO, the available videos in YouTube concerning Physical Activity during CoVid-19 lockdown only show a mean of 1.72 exercises, 2.83 for the high quality videos and no video showed more than 5 exercises.



Penrose Triangle

(“*impossibility in its purest form*”)

“The explanation requiring the fewest assumptions
is the most likely to be correct”
William of Ockham (Ockham´s Razor)

6

6. Conclusions

The following conclusions can be drawn from the information gathered from the data collection and the different statistical analyses carried out for these three studies that compose the research object of this thesis:

6.1. Article 1

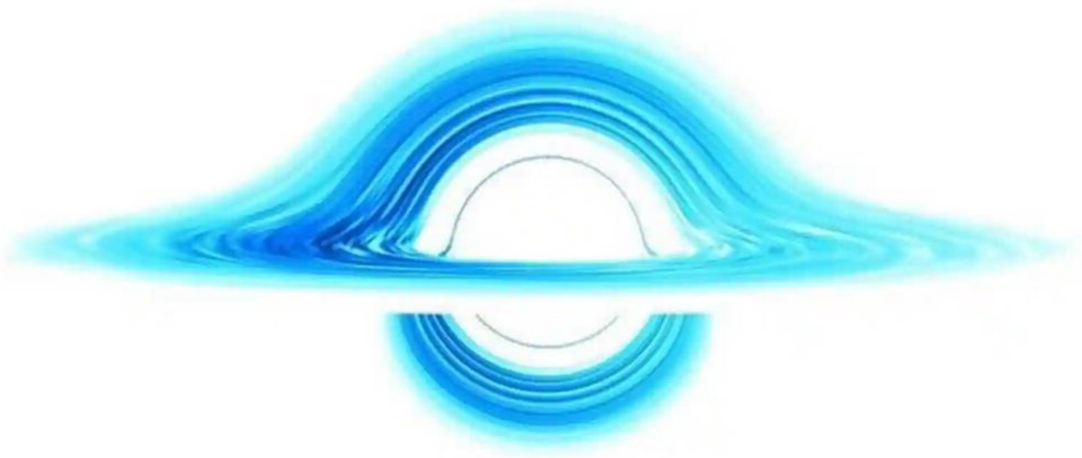
The quality of the information offered on YouTube about recommended postoperative shoulder exercises for Breast Cancer patients, according to the scores obtained in the quality scales, is "*High*".

6.2. Article 2

Article 2 found that the quality of the videos available on YouTube regarding the recommended exercises for pelvic floor for Prostate Cancer patients after prostatectomy surgery, according to the scores obtained in the quality scales, is "*High*".

6.3. Article 3

The quality of available videos in YouTube concerning Physical Activity during CoVid-19 lockdown is slightly lower than in the previous two articles and does not reflect WHO's recommendations. Nevertheless, according to quality scales, the quality of these videos is "*Average*".



Most realistic image of a black hole up to its time.

"Once we accept our limits,

we go beyond them"

Albert Einstein

7

7. Limitations and future perspectives

7.1. Limitations

Adhering to the principles of simulating what users see and in order to simulate standard user behaviour, no searches were conducted under incognito mode to avoid the influence of browsing history and geographical locations. Although YouTube content changes over time, our analysis represents the state of the videos at one specific time. Furthermore, the results of this study can only be considered for this website since it only takes this one into account, as mentioned in the selection criteria.

Only videos directly accessed on YouTube upon searching "*Exercises after breast cancer surgery*" in Article 1, "*Prostate cancer-exercises-pelvic floor*" for Article 2 and "*CoVid Exercises Home*" for Article 3 were included in this study. External links from other medical related websites were excluded in our analyses.

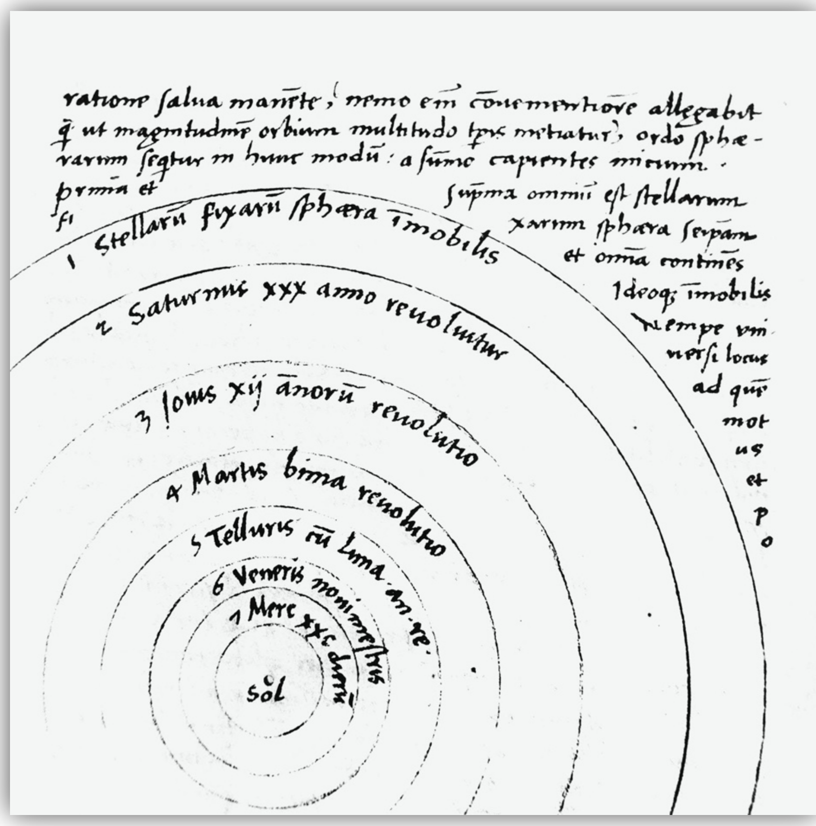
As a massive Internet-based platform, YouTube searches are constantly evolving as new videos are constantly being uploaded, viewed, and rated. Furthermore, because of this study tried to reproduce the typical user's search behaviour, our search was limited to the first 150 videos. As other studies have previously explained, our methodology was designed to replicate the average patient's search attempt, since most Internet users do not pay attention further than the first 50 search results.

7.2. Future perspectives

A verification and validation process of the information available on the web is considered necessary, as well as educational programs to facilitate people's access to the most reliable information (Ahmad et al., 2020; Bahar-Ozdemir et al., 2021; Fode et al., 2020; Warren, Sawhney, et al., 2021). Educational and health institutions, as well as health professionals, governmental health authorities and policymakers should be involved in the proper development of policies that improve the information available on the web with the aim of creating a positive impact on people's health-related behaviour, as recommended in other studies (Bahar-Ozdemir et al., 2021; Brar et al., 2021; Gulve et al., 2022; Kodonas & Fardi, 2021; Warren, Sawhney, et al., 2021). Effective strategies and policies capable of indicating the quality of this information are needed to filter out erroneous or non-rigorous information that may affect people's health. These tools should help any user/viewer to distinguish videos of high and low quality (Enver et al., 2020; Fode et al., 2020; Yalkin et al., 2022; X. Zhang et al., 2022). Due there is no mark to identify high-quality videos, it is recommended for their easy identification that patients attend to the source of production. Since Academic Institutions, NGOs, and/or Health Institutions show higher figures in quality scales, these should be the reference. High-quality videos were uploaded by academic sources and health professionals (Fode et al., 2020; Ozduran & Büyükçoban, 2022) and that is the reason why these groups should be aware of the situation and promote, produce and upload more high-quality videos (Crutchfield et al., 2021; Goobie et al., 2019; Güloğlu et al., 2022; Lang et al., 2022; Morra et al., 2022; Passos et al., 2020). Trusted institutions could use this social network to disseminate high quality information for patients and the public in general. It would also be advisable to discard those videos that contain commercials and could present any conflict of interest (Brar et al., 2021; Vu et al., 2021).

Supported by Machine Learning techniques and unsupervised learning algorithms, what is immediately plausible, and therefore should be planned for, is a federation of 'narrow' and 'targeted' machine learning systems capable of addressing core information processing problems across a health system, increasing the capabilities of human decision-makers and, in doing so, setting new standards of effectiveness and efficiency in clinical and managerial operations. Computer science can advance evidence-based medicine through development, testing, and refinement of artificial intelligence tools to deploy automation, creating 'living' evidence syntheses (Amezcua-Prieto et al., 2020; Bauer & Lizotte, 2021; Yoldemir, 2020). This is a major opportunity for health system transformation, as the cost of increasing decision-making capacity across the health system is not very high (Benke & Benke, 2018; Panch et al., 2018; Schwalbe & Wahl, 2020; Singh, 2019; Thiébaut & Thiessard, 2018).

Future studies and interventions, guided by the principles of translational medical research, are necessary to increase the quality of health information in YouTube videos. Through the development of effective policies that improve health behaviours of patients and users in general, reliable and good quality information must be differentiated, avoiding the inherent risks in searches for information related to health in social networks.



Fragment of the heliocentric image proposed by Copernicus published in his book "*De revolutionibus orbium coelestium*" in 1543, the year he died.

"If I have seen further it is by standing on the shoulders of Giants."

Issac Newton

8

8. Articles published for the thesis.

8.1. Article 1

"Quality Analysis of YouTube Videos Presenting Shoulder Exercises after Breast Cancer Surgery"

Citation:

Rodriguez Rodriguez A, M, Blanco-Diaz M, Lopez Diaz P, de la Fuente Costa M, Dueñas L, Escobio Prieto I, Calatayud J, Casaña J: Quality Analysis of YouTube Videos Presenting Shoulder Exercises after Breast Cancer Surgery. Breast Care 2022;17:188-198. doi: 10.1159/000518265

Research Article

Breast Care

Breast Care
DOI: 10.1159/000518265

Received: April 6, 2021
Accepted: July 2, 2021
Published online: August 25, 2021

Quality Analysis of YouTube Videos Presenting Shoulder Exercises after Breast Cancer Surgery

Alvaro Manuel Rodriguez Rodriguez[a] María Blanco-Diaz[a] Pedro Lopez Diaz[a] Marta de la Fuente Costa[a] Lirios Dueñas[b] Isabel Escobio Prieto[c] Joaquin Calatayud[b] Jose Casaña[b]

[a]Faculty of Medicine and Health Sciences, University of Oviedo, Physiotherapy and Traslational Research Group (FINTRA-RG), Institute of Health Research of the Principality of Asturias (ISPA), Oviedo, Spain; [b]Exercise Intervention for Health Research Group (EXINH-RG), Department of Physiotherapy, University of Valencia, Valencia, Spain; [c]Department of Physiotherapy, University of Sevilla, Seville, Spain

Keywords

YouTube · Breast cancer · Healthy behavior · Shoulder · Exercises · Public health

Abstract

Background: The prolonged immobilization suggested after breast cancer (BC) surgery causes morbidity. Patients search the Internet, especially social networks, for recommended exercises. ***Objective:*** The aim of this observational study was to assess the quality of YouTube videos, accessible for any patient, about exercises after BC surgery. ***Methods:*** A systematic search was performed on YouTube. One hundred and fifty videos were selected and analyzed. Two statistical analyses were conducted based on machine-learning techniques. Videos were classified as "Relevant" and "Non-Relevant" using principal component analysis models. Popularity was evaluated by Video Power Index (VPI), informational quality and accuracy were measured using the DISCERN Scale and Global Quality Scale (GQS). Scoring criteria for exercises were established according to the exercises recommended by the Oncology Section of the American Physical Therapy Association (APTA). Interobserver agreement and individual correlations were statistically examined. ***Results:*** DISCERN scored a mean of 50.97 (standard deviation [SD] 19.19). HONcode scored 78.30 (11.02) and GQS scored 3.49 (0.74). Average number of views was 53,963 (SD 67,376), mean duration was 9:42 min (9:15), mean days online was 2,158 (922), mean view ratio was 27.14 (30.24), mean likes was 245 (320.5), mean dislikes was 13.4 (14.2), and mean VPI was 93.48 (5.42). ***Conclusion:*** The quality of YouTube videos of recommended exercises post-BC surgery is high and can be a translational activity to improve patients' behavior. Health institutions and NGOs, with higher popularity levels than academic institutions, should consider this information when implementing new policies focused on video quality which can contribute to adaptive behavior in patients.

Introduction

According to Globocan 2020, breast cancer (BC) is the most common cancer overall with an incidence of 2,261,419 women in 2020 and exhibits higher incidence in developed countries [1]. There were over 2 million new cases in 2020 [2].

Women with BC are often treated with different types of surgery, including mastectomy and/or breast reconstruction [3].

In BC survivors, the prolonged immobilization caused by fear and/or pain, as well as type of surgery are factors in morbidities related to adhesive capsulitis, lymphedema, and shoulder pain [4, 5]. The weakness and limitation

karger@karger.com
www.karger.com/brc

Correspondence to:
María Blanco Díaz, blancomaria@uniovi.es

Karger

of patients' range of motion of the arm caused by pain and stiffness are considered one of the major postoperative complications of BC treatment [3] and are accompanied by the deterioration and decrease in upper limb function, which limits patients' ability to perform daily activities, such as dressing, bathing, and combing their hair. This, in turn, impairs their quality of life [4, 5].

Exercises during postoperative care of women with BC are important because it improves pain perception and contributes to enhancing quality of life. Early postoperative exercise is safe and improves shoulder function and general functional ability [4, 5]. Most BC survivors can safely participate in regular moderate-intensity exercise, which can improve several functional, physiological, and psychological parameters [6]. There is no consensus regarding how much and which exercises are best suited to promote postoperative shoulder functional recovery. Emphasis is on stretching the cervical region and active movement of the arm (shoulder flexion, extension, abduction, and rotation) to treat postoperative pain and impaired range of motion following treatment [4, 7]. The Oncology Section of the American Physical Therapy Association (APTA) [8], in collaboration with the American Cancer Society (ACS), recommends a series of the following seven exercises [9]: wand exercise (increases the ability to move shoulders forward), elbow winging (increases movement in the front of the chest and shoulder), shoulder blade stretch (increases shoulder blade movement), shoulder blade squeeze (increases shoulder blade movement and improves posture), side bends (increase movement of trunk and body), chest wall stretch (stretches the chest), and shoulder stretch (increases shoulder mobility). Strengthening exercises improve shoulder range of motion in flexion, abduction, and external rotation after surgery. Shoulder abduction and external rotation revealed less recovery, regardless of the intervention [10]. A home-based exercise program improved symptoms in the affected upper limb and led to improved quality of life in women [11]. Progressive upper limb exercises and muscle relaxation training positively affected upper limb function and health-related quality of life and can be used as an optional strategy [12].

Many patients search for medical information on the Internet, but lack the tools to evaluate the advice given [13]. As of March 3, 2020, an estimated 58.7% of the global population, i.e., 4,574,150,134 people, had access to the Internet [14]. YouTube is one of the most popular Internet sites and also operates as a primary Internet platform for consumer-targeted health information. Despite Internet use among BC patients being reported as 42–49%, the information obtained is not discussed with their physicians. According to Google Trends, a platform that assesses the popularity of online topics across the globe, BC is the most often searched cancer on YouTube [15].

Advances in eHealth technology have cultivated transactional opportunities for patients to access, share, and monitor health information [16]. Patients and health professionals often search the Internet for information regarding many health-related topics. It has been proven that 81% of all Internet users go online and search for information related to health [17]. Generally, patients search for detailed information about recommended exercises. Therefore, the authorship, quality, and validity of the information contained in the videos must be considered [18].

Although many health-related videos on YouTube are deemed educationally useful and high quality, some studies reveal that this is not always the case and some have commercial content designed to sell products or services, which may have serious implications for consumer attitudes and medical decision-making [19, 20].

Video-sharing websites must be understood as social media sites where there is no editorial selection or quality assessment. Likewise, potentially harmful and inaccurate information about science and biomedical topics can be disseminated. The kind of content stored on YouTube and the quality of the information are unclear [21]. A platform like YouTube has the potential to be an important resource for sharing and disseminating health-related information [22]. One question that has been highlighted by biomedical institutions is whether YouTube provides users and (potential) patients with accurate and helpful information or if the videos are possibly harmful and misleading. Healthcare professionals and organizations should be encouraged to provide more beneficial material and animated videos to people looking for comprehensive, reliable information on the Internet [23]. Healthcare providers and government agencies have expressed their concern about the veracity and quality of the information available on this platform, as the health-content videos uploaded by various sources can be misleading and present inaccurate information to patients [23].

Methods

Search Strategy

On March 31, 2019, a search was conducted on http://www.youtube.com using the following search term: "Exercises after breast cancer surgery."

The first 150 videos available to viewers were selected. We aimed to replicate a simple search strategy that could be conducted by any person. We did not restrict the search using filters; hence, YouTube sorted video results by their relevance according to the patented ranking algorithm active on that specific day. All the videos were added to a spreadsheet and then submitted for screening for duplicates, as well as to apply the inclusion and exclusion criteria by the research team. Exclusion criteria were non-English language, less than 5,000 views, duplicated videos, and/or related with advertisements. Video URLs were used to identify videos for subsequent screening and coding. Finally, 51 videos were assigned to two different examiners who viewed, analyzed, and evaluated them independently over a period of 5 weeks (Fig. 1).

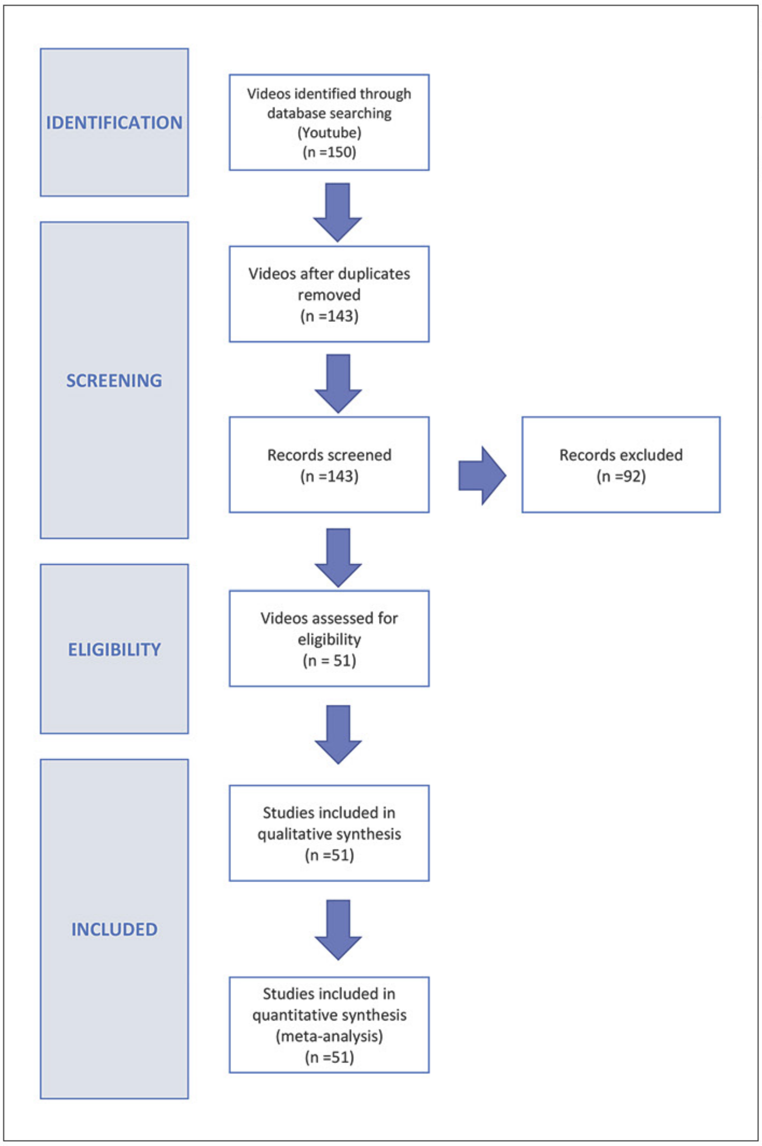


Fig. 1. Flow diagram for included videos. From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7):e1000097. doi:10.1371/journal.pmed1000097.

Outcome Measures

Based on their production source, the videos were categorized into five groups: health organization (clinic/hospital), individual, academic institution, media (TV, newspaper, etc.), and NGOs. Videos were also coded according to their continent of origin (America, Europe, Asia, Africa, Australia). Similarly, descriptive characteristics of each video (view counts, likes, dislikes, origin, days online, author, and duration) were collected. Video popularity was assessed using the Video Power Index (VPI) [(like count/(dislike count + like count)) × 100)] [18, 24, 25] and view ratio (views count/days online) [26]. The educational quality of the 51 selected videos was determined using the DISCERN instrument (Quality Criteria for Consumer Health Information) [27] and the Global Quality Scale (GQS) [28].

The DISCERN questionnaire was created by the Division of Public Health and Primary Care at Oxford University, London, to gauge the quality of information regarding treatment choices for health problems and was first published in 1999 [29]. The questionnaire consists of a total of 15 questions, in addition to an overall quality rating. Each question represents a different quality criterion and is rated from 1 to 5 (1: very poor, 2: poor, 3: average, 4: high, and 5: very high quality). These 16 questions are classified into three sections: reliability (questions 1–8), information quality on treatment choices (questions 9–15), and overall score (question 16). Thus, the maximum total score is 80 points; higher scores indicate higher quality [30].

GQS assesses the content quality of online resources. One point is assigned for each of the 5 identifiable criteria present in a video,

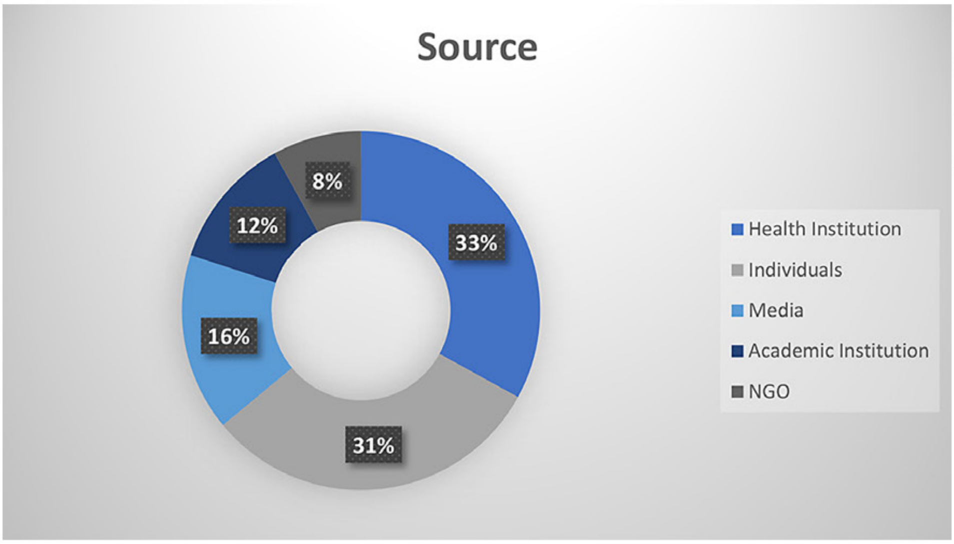


Fig. 2. Distribution of the videos based on source.

with 5 being the highest educational quality [28]. This scale covers accessibility and quality of the information, the overall flow of information, and how useful it would be for a user [26, 31].

Each video was also appraised using the HONcode Health Evaluation Tool [32, 33]. The Health On the Net Foundation (HON), a non-profit organization under the auspices of the United Nations, has elaborated the code of conduct to help standardize the reliability of medical and health information available on the World Wide Web [33, 34]. HONcode is the most widely used quality evaluation tool [35]. HON certification is subject to an annual review by the HON Foundation, which also responds to any reported violations by Internet users. HONcode does not rate the quality of the information provided by a website, it defines a set of rules to hold website developers to basic ethical standards with respect to the information they present. It also helps viewers to be aware of the source and the purpose of the data they are viewing. Determining a set of reliability and credibility parameters for the information available on health websites based on 8 principles, it evaluates the reliability of web pages measured from 0 to 100% [34].

The scoring criteria for exercises were set according to recommendations by the Oncology Section of the APTA in collaboration with the ACS [3]. Since 7 exercises are recommended, the scoring system was based on a scale of 0 to 7 (1 point for each exercise coinciding with the 7 recommended by APTA).

Statistical Analysis

Two different analyses were performed on the dataset: the first analysis was based on machine learning techniques, considering a binary classification problem, and thereby dividing the sample into "Relevant Videos" (C1) and "Not Relevant Videos" (C2). These two classes were defined in three cases, using the means of DISCERN, number of exercises, and Principal Component Analysis (PCA) [36]. PCA visualizes the information in a dataset described by multiple, interrelated variables. The information in a given dataset corresponds to the total variation. PCA identifies directions (or principal components) along which there is maximum variance in the data. It is used to extract the important information from a multivariate dataset and express this information as a set of few new variables, i.e., the principal components. Thus, PCA reduces the dimensionality of multivariate data to two or three principal components (in this case, two principal components were used to group data) that can be visualized graphically, with minimal loss of information.

The objective of this first analysis was to determine those variables which best distinguish these two video classes. The classification problem was studied in three cases, when both classes were defined by the following variables:

1. DISCERN: In this case, samples in the first class C1 (relevant videos) were considered those videos whose DISCERN variable value was greater than the mean for this variable; the rest of the videos were defined as C2.
2. Number of exercises: Similar to the previous variable.
3. PCA: This is a binary variable that already contains the class of each sample after performing a PCA and a k-means algorithm. In this case, the sample grouping into C1 and C2 is given by the variability in data and there is no "a priori" information as to the relevance of the videos in each class. Consequently, reducing the number of variables to two (as per PCA) that group the most important information of all the initial variables means that no information gets lost.

For each of the three cases, a classification algorithm was applied to attain a minimum list of variables that best assigns the samples into two groups. Accuracies are determined by leave-one-out cross validation using a nearest-neighbour classifier. The discriminatory power of the variables was established according to their Fisher's ratio (FR). Variables with an elevated FR are highly discriminatory, inasmuch as they have low intraclass dispersion and high interclass distance. In a binary classification, the FR of the variable j is given by:

$$\frac{\left(\mu_{j1} - \mu_{j2}\right)^2}{{\sigma_{j1}}^2 + {\sigma_{j2}}^2},$$

where μ_{ji} is a measure of the center of mass of the probability distribution of the variable j in class i ($i = 1,2$) and σ_{ji} is a measure of its dispersion within this class. Variables with a high FR are very discriminatory, since they have a low intraclass dispersion and high interclass distance.

Table 1. Descriptive statistics

	Mean ± standard deviation			Median (min–max)		
	observer 1	observer 2	observers' mean	observer 1	observer 2	observers' mean
Video length, min	9:42±9:15			6:40 (0:54–11:11)		
Views	53,963±67,376			24,595 (6,024–330,608)		
Days online	2,158±922			2,175 (377–4,045)		
View ratio (views/days)	27.14±30.24			14.85 (1.84–155.96)		
Likes	245.0±320.5			139 (2–1,384)		
Dislikes	13.4±14.2			8 (0–54)		
Likes/day	0.151±0.222			0.061 (0.001–0.962)		
Dislikes/day	0.008±0.013			0.004 (0.000–0.080)		
VPI, %	93.48±5.42			94.87 (74.19–100.00)		
DISCERN score	49.80±9.94	52.14±11.59	50.97±10.19	51 (27–68)	53 (24–70)	53.0 (25.5–68.0)
HON score	78.06±11.85	78.55±11.31	78.30±11.02	81 (52–95)	82 (53–94)	82.0 (56.0–93.5)
GQS score	3.57±0.70	3.41±0.92	3.49±0.74	4 (2–5)	4 (1–5)	3.5 (1.5–5.0)

In addition to FR, Pearson, Kendal, and Spearman correlation factors of the variables with the defined classes were computed. These factors also reveal the relevance (discriminatory power) of the variables for the classification criteria.

The second analysis was a statistical one using a *t* test and the Wilcoxon test, illustrating how significant the differences between both classes are from the perspective of each variable. The statistical analysis also shows the relevance of each variable regarding the grouping of the samples into C1 and C2, in each of the three cases. The *t* test determines if the variable's means in both datasets are significantly different from each other. The null hypothesis H0 assumes that there is no difference between groups. When the statistical analysis yields 1, it means that there is enough evidence to reject said hypothesis and if it results in 0, the null hypothesis must be accepted. The accepted value level of significance or error probability is alpha <0.05. Consequently, the Wilcoxon test essentially calculates the difference between each set of pairs and analyzes these differences, whereas the null hypothesis H0 refers to the equality of population medians of two groups of samples (C1 and C2).

Results

Video Characteristics

Most of the videos (33%) were produced by health institutions, followed by individuals (31%), media sources (16%), academic institutions (12%), and NGOs (8%) (Fig. 2).

The continent of origin was America in 41% of the cases, Europe (14%), Australia (4%), and Asia (2%).

Descriptive Statistics

The mean number of views was 53,963 (standard deviation [SD] 67,376), mean video duration was 9:42 min (SD = 9:15), mean number of days online was 2,158 (SD = 922), mean view ratio count was 27.14 (SD = 30.24), mean likes count was 245 (SD = 320.5), mean dislikes count was 13.4 (SD = 14.2), and mean VPI was 93.48 (SD = 5.42). Table 1 displays the basic data examined by descriptive statistics for the videos included.

An intraclass correlation coefficient (ICC) analysis was conducted to gauge inter-examiner concordance with 95% confidence intervals (CI) based on mean rating ($k = 2$), consistency, two-way random model, and Pearson's correlation method. The level of significance was set at $p < 0.05$. Inter-rater agreement scored 0.9860 in this study.

The average obtained through both independent assessments was used for each video. DISCERN, HONcode, and GQS scores of both observers were averaged to calculate the mean scores for each video. Considering average results, the mean DISCERN score was 50.97 (SD = 19.19) with a median of 53.0 (minimum 25.5 and maximum 68.0), the mean HONcode score was 78.30 (SD = 11.02) with a median of 82.0 (minimum 56.0 and maximum 93.5), and the mean GQS score was 3.49 (SD = 0.74) with a median of 3.5 (minimum 1.5 and maximum 5.0). The basic data examined by descriptive statistics for DISCERN, HONcode, and GQS results can be seen in Table 1.

Table 2 shows how videos performed for each DISCERN question and presents the average scores (1–5) and SD. The median with the minimum and maximum are also displayed. The question about "whether the aims are clear" (Question 1) scored highest, followed by the question about "achieving the aims" (Question 2). The lowest-scoring questions were "whether it refers to areas of uncertainty" (Question 8) and "if no treatment was used" (Question 12).

According to the mean DISCERN scores of both observers, the quality of the videos was deemed very poor in 2% of the cases, poor in 10%, average in 22%, high in 59%, and very high in 8% (Fig. 3). However, the mean GQS

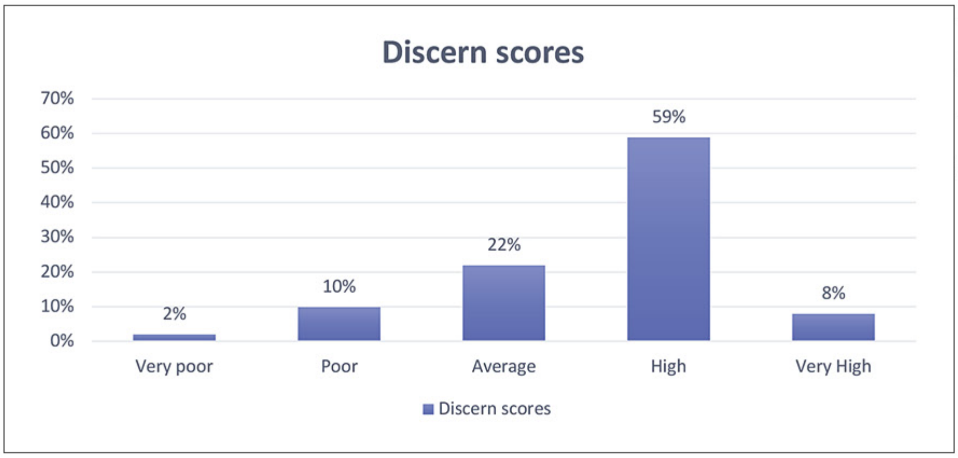


Fig. 3. DISCERN scores.

Table 2. Average score per DISCERN question

	DISCERN questions	Mean score ± SD	Median (min–max)
1	Are the aims clear?	4.50±0.91	5.0 (1.0–5.0)
2	Does it achieve its aims?	4.26±0.83	4.5 (1.5–5.0)
3	Is it relevant?	4.15±1.06	4.5 (1.0–5.0)
4	Is it clear what sources of information were used to compile the publication (other than the author or producer)?	2.20±0.99	2.0 (1.0–4.0)
5	Is it clear when the information used or reported in the publication was produced?	2.62±1.05	3.0 (1.0–4.0)
6	Is it balanced and unbiased?	2.65±1.02	2.5 (1.0–4.0)
7	Does it provide details of additional sources of support and information?	3.25±1.05	3.0 (1.0–5.0)
8	Does it refer to areas of uncertainty?	1.87±0.85	1.5 (1.0–3.5)
9	Does it describe how each treatment works?	3.25±1.06	3.5 (1.0–5.0)
10	Does it describe the benefits of each treatment?	3.44±1.02	3.5 (1.5–5.0)
11	Does it describe the risks of each treatment?	2.62±1.00	2.5 (1.0–5.0)
12	Does it describe what would happen if no treatment is used?	2.07±0.84	2.0 (1.0–4.0)
13	Does it describe how the treatment choices affect overall quality of life?	3.10±1.00	3.0 (1.0–5.0)
14	Is it clear that there may be more than one possible treatment choice?	3.41±1.03	3.5 (1.0–5.0)
15	Does it provide support for shared decision-making?	3.95±0.84	4.0 (1.5–5.0)
16	Based on the answers to all of the above questions, rate the overall quality of the publication as a source of information about treatment choices	3.64±0.81	4.0 (1.5–5.0)

Table 3. Pearson correlation coefficient of every variable in respect to DISCERN

Variable	Pearson coefficient	*p* value
HONcode	0.70	<0.001
GQS score	0.65	<0.001
Likes/day	0.31	0.0282
VPI	–0.28	0.0435
Views/day	0.27	0.0574

scores showed the quality of the videos to be very poor in 6% of the cases, poor in 6%, average in 41%, high in 45%, and very high in 2% (Fig. 4).

Similarly, Pearson's correlation coefficients for each DISCERN variable were computed. These factors also point to the relevance (discriminatory power) of the remaining variables with a significance level of $p < 0.05$ (Table 3).

As for the relation between the production source and educational quality of the videos, according to DISCERN, the videos made by academic institutions scored higher (56.25), followed by health institutions (54.85), NGOs (51.38), media (50.94), and, finally, those created

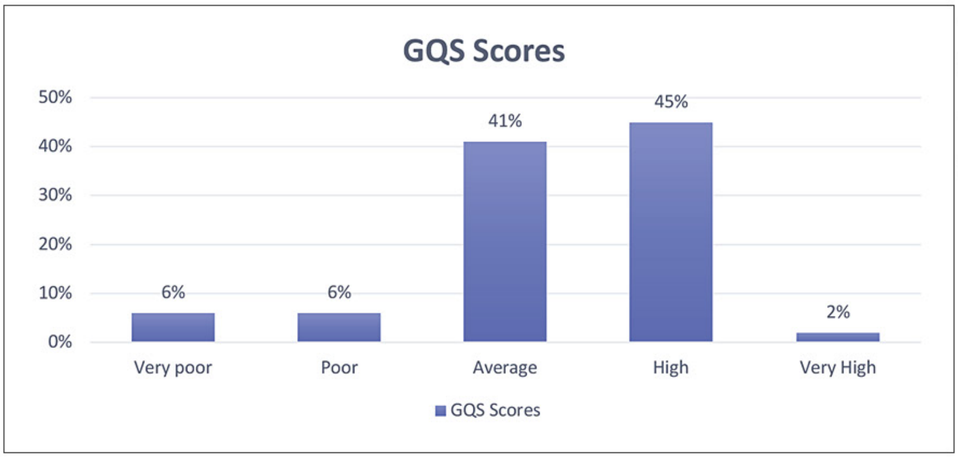


Fig. 4. GQS scores.

Table 4. Information sources, origin, and educational quality

		DISCERN, mean	GQS, mean	View ratio, mean
Source	Academic institutions	56.25	3.83	62.96
	Health institutions	54.85	3.74	24.95
	NGOs	51.38	3.63	28.78
	Media	50.94	3.31	10.40
	Individuals	44.78	3.16	26.24
	Total	**50.97**	**3.49**	**27.14**
Origin	America	51.73	3.52	23.75
	Europe	51.57	3.79	45.03
	Asia	47.00	3.50	50.69
	Australia	35.25	1.75	20.61
	Total	**50.97**	**3.49**	**27.14**

by individuals (44.78). Likewise, the GQS, the other variable that alludes to educational quality, ranks the origins in the same order: academic institutions (3.83), followed by health institutions (3.74), NGOs (3.63), media (3.31), and, finally, individuals (3.16). These data and view ratios can be seen in Table 4. The same table shows that higher scores on the DISCERN scale were given to videos from America (51.73), followed Europe (51.57), Asia (47.00), and Australia (35.25). However, the GQS scale gave higher scores to videos made in Europe (3.79), followed by America, (3.52), Asia (3.50), and, finally, Australia (1.75).

VPI mean scores for each of the five categories in which the DISCERN variable classified the videos were also analyzed. Videos with very poor DISCERN scores had a mean VPI of 100, videos classified as poor by DISCERN had a VPI of 84.78, average videos had a VPI of

Table 5. Mean number of exercises according to GQS and DISCERN classifications

	Quality	Mean number of exercises
GQS	1 – very poor	0.67
	2 – poor	1.33
	3 – average	1.10
	4 – high	2.30
	5 – very high	7.00
	Total	**1.75**
DISCERN	1 – very poor	0.00
	2 – poor	1.20
	3 – average	1.45
	4 – high	2.10
	5 – very high	1.00
	Total	**1.75**

Table 6. *p* values of GQS, DISCERN, HONcode, and number of exercises regarding PCA according to *t* test and Wilcoxon test

Variable	*t* test		Wilcoxon test	
	H	*p*	H	*p*
GQS	1	0.0000	1	0.0000
Exercises	1	0.0000	1	0.0000
DISCERN	1	0.0007	1	0.0015
HONcode	1	0.0043	1	0.0019

Table 7. ICC intervals

Interval	Reliability*
Less than 0.5	Poor
0.5–0.75	Moderate
0.75–0.90	Good
Greater than 0.90	Excellent

* Based on the 95% confidence interval.

94.66, videos scored as high for DISCERN had a VPI of 94.85, and videos classified as very high as per DISCERN had a mean VPI of 89.55.

Of all the videos included, the mean number of APTA-recommended exercises was 1.75 (SD 1.91). As to GQS classification, "very high" videos included a mean of 7, "high" contained a mean of 2.3, "average" videos had 1.10, "poor" included 1.33, and "very poor" had a mean of 0.67 APTA-recommended exercises. The videos classified by DISCERN as "very high" included a mean of 1 exercise, "high" videos contained a mean of 2.10 exercises, "average" videos had 1.45 exercises, "poor" videos covered 1.20, and "very poor", 0.00 APTA-recommended exercises (Table 5).

The analysis of the number of APTA-recommended exercises in each video indicates that 45% of the videos included no such exercise. Six percent of the videos showed only one such exercise, 16% included 2, 6% demonstrated 3, 22% had 4, 4% contained 5, and 2% of the videos covered 7 APTA-recommended exercises.

Videos with the highest mean VPI (96.87) showed 5 of the recommended exercises and are also the ones that present the highest view ratio (44.36) and the highest likes ratio (0.3482).

In the second statistical analysis (Table 6), *p* values for GQS, DISCERN, HONcode, and number of exercises regarding PCA are <0.01, indicating a high, statistically significant relation.

Discussion

This study was undertaken to better comprehend the nature of evidence independently accessed by patients on a massive online, media-sharing platform. We found that, although variable in source and content, the quality of the information offered on YouTube about recommended postoperative shoulder exercises for BC patients is high.

The increasing use of new technologies and vast amount of content available have turned the Internet into an important source of health information, and both health professionals (who use new technologies more frequently to communicate with patients and achieve the necessary adherence and interaction in health processes) and patients (who use it as a source of information) are at risk inasmuch as the information available on the web is not subject to any kind of inspection. Consequently, this can lead to erroneous and even harmful messages being conveyed to users and/or viewers.

In the search process to improve health processes, rigorous education and dissemination are necessary, preferably carried out by health professionals. In the case of the videos analyzed, most (33%) were produced by health institutions, which can be contemplated as a kind of a guarantee, given that they ranked second on the quality scales. This percentage is very similar to the number of videos made by individuals, which, though extremely numerous, obtained the lowest scores on the educational quality scales. Unfortunately, the percentage of videos made by academic institutions was much lower (12%), which agrees with other studies that state that health institutions are underrepresented in the publication of medical information videos [24]. This is a shame because they could consider academic projects to increase their presence and flood the Internet with good-quality information since, in light of the results obtained, the quality of their videos was the best of the entire sample analyzed. As in similar studies, it was often difficult to ascertain who developed the video content, as well as for viewers to assess the reputation of the sources [37].

The average length of the videos in our study was 9:42 min, which is consistent with previous studies that have reported an average duration of 6:17–10:35 min [24, 38]. No statistically significant relationship was found between video length and their quality or popularity, contrary to the study by Aydin and Akyol [25], who found that the videos with the longest duration and the highest VPI appeared to be associated with higher quality scores.

ICC is a widely used test-retest, intra-rater, and inter-rater reliability index. Based on the 95% CI of the ICC estimation, the inter-reviewer agreement for this study was 0.9860, pointing toward excellent concordance between both examiners, according to the ICC intervals shown in Table 7 [39].

Codes like HONcode do not assess the accuracy of medical information but refer to the ethics of the infor-

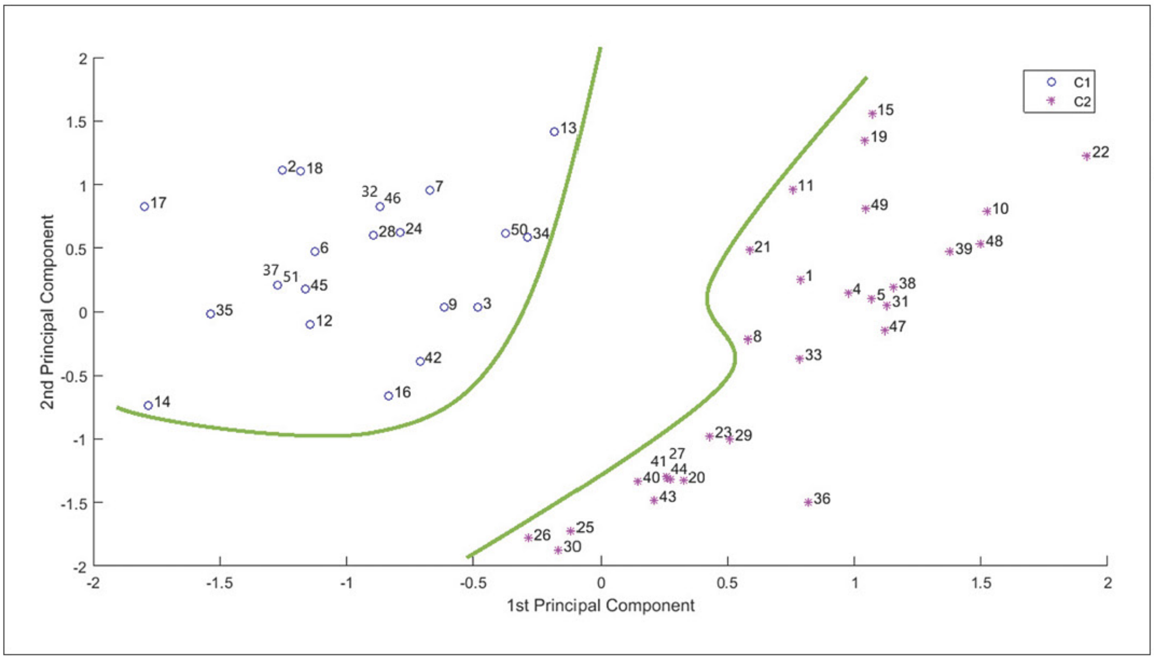


Fig. 5. PCA components diagram.

mation presented. In this study, HONcode yielded high scores (78.30 ± 11.02). Multiple studies have found that HONcode certification was associated with better quality of information and greater reliability [35], in line with this study, which shows a statistically significant relation between DISCERN and HONcode scales (Pearson's coefficient = 0.70; $p < 0.001$).

A statistically significant relation between DISCERN and GQS scales was also detected (Pearson's coefficient = 0.65; $p < 0.001$). Analyzing the scores obtained to appraise the educational quality of videos with both scales, it should be noted that in both cases, most videos were high or very high quality (67% and 47% of the cases, respectively), average quality was found in 22% and 41%, and very few videos had poor or very poor educational quality (12% in both cases). This contrasts with earlier studies that have identified a worrisome amount of poor-quality and erroneous health-related information in YouTube videos [19, 25, 26], although these studies did not deal with BC.

In the analysis performed, videos produced by a health institution, despite not being the most viewed, once viewed, they were found to receive many user interactions, placing them in second place in interactions, and to receive the second most "likes." These videos were produced in Europe and America and scored high, behind videos created by academic institutions, as per DISCERN and GQS scales.

Videos produced by the media (newspapers, TV, etc.) are the least viewed and receive fewer "likes" and "dislikes." All originate in America and have average educational quality. Videos created by NGOs are in second position, behind those produced by academic institutions, in number of views. They received the most likes and the second most dislikes. Despite their high number of daily views, their educational quality is low (mean 3.63 [SD 0.48] in GQS and 51.38 [SD 7.25] in DISCERN).

It is important to note that the videos with the highest educational quality in this study are produced by academic institutions, which are also the most viewed (views per day), but whose popularity, as measured by VPI, is the lowest because, though they receive many likes, they receive the most dislikes. There is a statistically significant relation ($p = 0.0435$) with a negative Pearson's coefficient (–0.28) between DISCERN and VPI variables (Table 3). This result suggests that the quality and content of the videos uploaded by academic institutions may not be sufficient to connect with the user (a task that is achieved by those produced by health institutions). This is consistent with other studies that reflect that high-quality videos are not the most popular [24]. In contrast, poorer educational quality videos made by individuals are more popular than academic videos, consistent with the study by Radonjic et al. [40], who evidenced how educational videos made by nondoctors were significantly more popular than those made by doctors.

As regards the PCA perspective from the first statistical analysis, if the coordinates of each video are graphically represented in a coordinate system in which the axes refer to the principal components, videos are naturally grouped into two groups of points, without any manual intervention or possible bias. These two groups, which can be seen in Figure 5, collect the videos with higher ratings in HONcode, DISCERN, GQS, and number of exercises in one group (C1) and those with lower ratings in another group (C2) (Fig. 5).

The statistical power of these two combined analyses indicates that the PCA analysis allows the videos to classify themselves into two groups (C1, high-quality videos and C2, low-quality videos) without any information being lost on any of the variables considered in the entire study. Thus, the principal components of the PCAs show a high statistical relationship with the variables that evaluate the quality of the videos, which enables the quality of the videos to be analyzed through the two principal components, instead of it having to be done through all the variables of the study. This is a huge advantage when carrying out this kind of analysis, considering that any important information is not lost through these two principal components because their level of relationship with the main variables that evaluate the quality of the videos is maximum and p values reveal a high, statistically significant relation.

The statistical power of these combined analyses is such that it allows the videos classification, according to their quality, considering two variables (principal components) that combine together the information of all the other variables analyzed in the study.

As other authors have already confirmed, the results presented in this study indicate that YouTube is a solid and useful source of information but, in general, the video platform is an inadequate source of information [25, 41] and different ways are needed to ensure that users access the most appropriate ones.

Bias

Adhering to the principles of simulating what users see, no searches were conducted under incognito mode to avoid the influence of browsing history and geographical locations.

Although YouTube content changes over time, our analysis represents the state of the videos at one specific time [40].

Only videos directly accessed on YouTube upon searching "Exercises after breast cancer surgery" were included in this study. External links from other medical-related websites were excluded in our analyses.

As a massive Internet-based platform, YouTube searches are constantly evolving as new videos are constantly being uploaded, viewed, and rated. Furthermore, our search was limited to the first 150 videos. As other studies have previously explained, our methodology was designed to replicate the average patient's search attempt, since most Internet users do not look past the first 50 search results [40].

Conclusions

Digital behavior change interventions have the potential to improve healthy behavior after BC surgery. YouTube is one of the most important development tools of the eHealth era. Videos are a simple, attractive, and affordable source of information, with vast dissemination capacity available. The quality of videos available on YouTube about recommended postoperative shoulder exercises for BC patients is high, according to DISCERN and GQS scores.

A verification and validation process of the information available on the web is considered necessary, as well as educational programs to facilitate people's access to the most reliable information. Educational and health institutions, as well as health professionals, government health authorities, and policymakers should be involved in the proper development of policies that improve the information available on the web with the aim of creating a positive impact on people's health-related behavior.

Future studies, guided by the principles of translational medical research, are needed to increase the quality of health information in YouTube videos as effective policies to improve health behaviors in patients after BC surgery.

Statement of Ethics

This study does not involve human participants and informed consent was therefore not required. This article does not contain any studies with animals performed by any of the authors.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contributions

Conceptualization: María Blanco Díaz, Alvaro Manuel Rodriguez. Data curation: Alvaro Manuel Rodriguez. Formal analysis: Alvaro Manuel Rodriguez Rodríguez, María Blanco Díaz. Original draft preparation: Alvaro Manuel Rodriguez Rodríguez. Investigation: Pedro López Díaz, Marta de la Fuente Costa, Isabel Escobio. Methodology: Jose Casaña Granell. Project administration: Lirios Dueñas. Resources: Joaquin Calatayud, Isabel Escobio. Supervision: Jose Casaña Granell, Joaquin Calatayud. Validation: María Blanco Díaz. Visualization: Pedro López Díaz. Writing: Alvaro Manuel Rodriguez Rodríguez, María Blanco Díaz. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

All data generated or analyzed during this study are included in this article. Further enquiries can be directed to the corresponding author.

References

1 Zhuang S, Li L, Zang Y, Li G, Wang F. RRM2 elicits the metastatic potential of breast cancer cells by regulating cell invasion, migration and VEGF expression via the PI3K/AKT signaling. Oncol Lett. 2020;19:3349–55.
2 The Global Cancer Observatory. All cancers 2020. https://gco.iarc.fr/today (accessed June 23, 2021).
3 Cancer Research UK. Exercises after breast cancer surgery. Am Cancer Soc. 2020.
4 Reigle BS, Zhang B. Women's Rehabilitation Experiences Following Breast Cancer Surgery. Rehabil Nurs. 2018;43:195–200.
5 Richmond H, Lait C, Srikesavan C, Williamson E, Moser J, Newman M, et al. Development of an exercise intervention for the prevention of musculoskeletal shoulder problems after breast cancer treatment: the prevention of shoulder problems trial (UK PROSPER). BMC Health Serv Res. 2018;18:463–12.
6 Adraskela K, Veisaki E, Koutsilieris M, Philippou A. Physical Exercise Positively Influences Breast Cancer Evolution. Clin Breast Cancer. 2017;17:408–17.
7 De Groef A, Van Kampen M, Dieltjens E, Christiaens MR, Neven P, Geraerts I, et al. Effectiveness of postoperative physical therapy for upper-limb impairments after breast cancer treatment: A systematic review. Arch Phys Med Rehabil. 2015;96:1140–53.
8 Worthen A. APTA. American Physical Therapy Association. n.d. https://www.apta.org/AboutUs/ (accessed May 12, 2020).
9 American Cancer Society. Breast Cancer: Treating Breast Cancer. Am Cancer Soc. 2019:1–120.
10 Ribeiro IL, Moreira RFC, Ferrari AV, Alburquerque-Sendín F, Camargo PR, Salvini TF. Effectiveness of early rehabilitation on range of motion, muscle strength and arm function after breast cancer surgery: a systematic review of randomized controlled trials. Clin Rehabil. 2019;33:1876–86.
11 Gautam AP, Maiya AG, Vidyasagar MS. Effect of home-based exercise program on lymphedema and quality of life in female postmastectomy patients: Pre-post intervention study. J Rehabil Res Dev. 2011;48:1261–8.
12 Zhou K, Wang W, An J, Li M, Li J, Li X. Effects of Progressive Upper Limb Exercises and Muscle Relaxation Training on Upper Limb Function and Health-Related Quality of Life Following Surgery in Women with Breast Cancer: A Clinical Randomized Controlled Trial. Ann Surg Oncol. 2019;26:2156–65.
13 Drozd B, Couvillon E, Suarez A. Medical YouTube videos and methods of evaluation: Literature review. JMIR Med Educ. 2018;4:e3.
14 World Internet Users Statistics and 2020 World Population Stats 2020.
15 Ayoub G, Chalhoub E. Evaluation of Breast Cancer Videos on Youtube in the Arab World. n.d.: 1–16.
16 Paige SR, Stellefson M, Krieger JL, Anderson-Lewis C, Cheong J, Stopka C. Proposing a Transactional Model of eHealth Literacy: Concept Analysis. J Med Internet Res. 2018;20:e10175.
17 Omnicore Agency. YouTube by the Numbers: Stats, Demographics & Fun Facts. OmnicoreagencyCom 2020. https://www.omnicoreagency.com/youtube-statistics/ (accessed April 13, 2020).
18 Atci AG. Quality and reliability of the information on YouTube Videos about Botox injection on spasticity. roneuro. 2019:473–7.
19 Loeb S, Sengupta S, Butaney M, Macaluso JN, Czarniecki SW, Robbins R, et al. Dissemination of Misinformative and Biased Information about Prostate Cancer on YouTube. Eur Urol. 2019;75:564–7.
20 Langford A, Loeb S. Perceived patient-provider communication quality and sociodemographic factors associated with watching health-related videos on YouTube: A cross-sectional analysis. J Med Internet Res. 2019;21:e13512.
21 Hartley J. Digital Futures for Cultural and Media Studies. 2012.
22 Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: A systematic review. Health Informatics J. 2015;21:173–94.
23 Aydin MF, Aydin MA. Quality and reliability of information available on YouTube and Google pertaining gastroesophageal reflux disease. Int J Med Inform. 2020;137:104107.
24 Kuru T, Erken HY. Evaluation of the Quality and Reliability of YouTube Videos on Rotator Cuff Tears. Cureus. 2020;12:e6852.
25 Aydin MA, Akyol H. Quality of Information Available on YouTube Videos Pertaining to Thyroid Cancer. J Cancer Educ. 2019;18:1–14.
26 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.
27 Charnock D. Discern Handbook; 1998.
28 Kunze KN, Cohn MR, Wakefield C, Hamati F, Laprade RF, Forsythe B, et al. YouTube as a Source of Information About the Posterior Cruciate Ligament: A Content-Quality and Reliability Analysis. Arthrosc Sports Med Rehabil. 2019;1:e109–14.
29 Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health. 1999;53:105–11.
30 Kılınç DD, Sayar G. Assessment of reliability of YouTube videos on orthodontics. Turkish J Orthod. 2019;32:145–50.
31 Esen E, Aslan M, Sonbahar BÇ, Kerimoğlu RS. YouTube English videos as a source of information on breast self-examination. Breast Cancer Res Treat. 2019;173:629–35.
32 HONcode: E-Guide for Health Consumers. n.d. https://www.hon.ch/HONcode/Pro/HealthEvaluationTool.html (accessed April 26, 2020).
33 Boyer C, Selby M, Scherrer JR, Appel RD. The health on the Net Code of Conduct for medical and health Websites. Comput Biol Med. 1998;28:603–10.
34 Boyer C. When the quality of health information matters: Health on the Net in the quality Standard for Information You can Trust; 2013.
35 Fahy E, Hardikar R, Fox A, Mackay S. Quality of patient health information on the internet: Reviewing a complex and evolving landscape. Australas Med J. 2014;7:24–8.
36 Zhang Z, Castelló A. Principal components analysis in clinical studies. Ann Transl Med. 2017;5:351–.
37 Heathcote LC, Pate JW, Park AL, Leake HB, Moseley GL, Kronman CA, et al. Pain neuroscience education on YouTube. PeerJ. 2019;7:e6603.
38 Gokcen HB, Gumussuyu G. A Quality Analysis of Disc Herniation Videos on YouTube. World Neurosurg. 2019;124:799–804.
39 Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15:155–63.
40 Radonjic A, Fat Hing NN, Harlock J, Naji F. YouTube as a source of patient information for abdominal aortic aneurysms. J Vasc Surg. 2020;71:637–44.
41 Steinberg PL, Wason S, Stern JM, Deters L, Kowal B, Seigne J. YouTube as Source of Prostate Cancer Information. Urology. 2010;75:619–22.

8.2. Article 2

“Quality Analysis of YouTube Videos Presenting Pelvic Floor Exercises after Prostatectomy Surgery”

Citation:

Rodriguez-Rodriguez, A.M.; Blanco-Diaz, M.; Lopez-Diaz, P.; de la Fuente-Costa, M.; Sousa-Fraguas, M.C.; Escobio-Prieto, I.; Casaña, J. Quality Analysis of YouTube Videos Presenting Pelvic Floor Exercises after Prostatectomy Surgery. J. Pers. Med. 2021, 11, 920. https://doi.org/10.3390/jpm11090920

Journal of Personalized Medicine

MDPI

Article

Quality Analysis of YouTube Videos Presenting Pelvic Floor Exercises after Prostatectomy Surgery

Alvaro Manuel Rodriguez-Rodriguez [1], Maria Blanco-Diaz [2,*], Pedro Lopez-Diaz [2], Marta de la Fuente-Costa [2], Maria Cruz Sousa-Fraguas [2], Isabel Escobio-Prieto [3] and Jose Casaña [1]

[1] Exercise Intervention for Health Research Group (EXINH-RG), Department of Physiotherapy, University of Valencia, 46010 Valencia, Spain; alvaro.manuel.rodriguez@gmail.com (A.M.R.-R.); jose.casana@uv.es (J.C.)
[2] Physiotherapy and Translational Research Group (FINTRA-RG), Institute of Health Research of the Principality of Asturias (ISPA), Faculty of Medicine and Health Sciences, University of Oviedo, 33006 Oviedo, Spain; lopezdpedro@uniovi.es (P.L.-D.); fuentemarta@uniovi.es (M.d.l.F.-C.); sousamaria@uniovi.es (M.C.S.-F.)
[3] Departamento de Fisioterapia, Facultad de Enfermería, Fisioterapia y Podología, Universidad de Sevilla, 41009 Sevilla, Spain; iescobio@us.es
* Correspondence: blancomaria@uniovi.es; Tel.: +34-667681858

Citation: Rodriguez-Rodriguez, A.M.; Blanco-Diaz, M.; Lopez-Diaz, P.; de la Fuente-Costa, M.; Sousa-Fraguas, M.C.; Escobio-Prieto, I.; Casaña, J. Quality Analysis of YouTube Videos Presenting Pelvic Floor Exercises after Prostatectomy Surgery. *J. Pers. Med.* **2021**, *11*, 920. https://doi.org/10.3390/jpm11090920

Academic Editor: Pierluigi Maria Rinaldi

Received: 22 July 2021
Accepted: 13 September 2021
Published: 15 September 2021

Abstract: Background: Prostate cancer (PC) is a major cause of disease and mortality among men. Surgical treatment involving the removal of the prostate may result in temporary or permanent erectile dysfunction (ED) and urinary incontinence (UI), with considerable impact on quality of life. Pelvic floor muscle training (PFMT) is one of the recommended techniques for the prevention, treatment, and rehabilitation of postoperative complications. The aim of this observational study was to assess the quality of YouTube videos—accessible to any patient—related to exercises after prostatectomy surgery. Methods: A systematic search was performed on YouTube on 24 September 2020. One hundred and fifty videos were selected and analyzed. Two statistical analyses were conducted based on machine-learning techniques, and videos were classified as 'Relevant' or 'Non-Relevant' using Principal Component Analysis (PCA) models. Two reviewers conducted independent analyses. Inter-observer agreement and individual correlations of video data were evaluated with the Intraclass Correlation Coefficient (ICC). Information quality, reliability, and accuracy were measured using the DISCERN Scale and Global Quality Score (GQS), while video popularity was evaluated using the Video Power Index (VPI). Results: DISCERN scored a mean of 3.35 and GQS scored 3.38. Average number of views was 124,354, mean duration was 14:42 min, mean days online was 1777, mean view ratio was 138.30, mean Likes was 1082, mean Dislikes was 68.58, and mean VPI was 92.28. Conclusions: The quality of the videos available on YouTube regarding the recommended pelvic floor exercises in PC surgery, according to the scores obtained, is High. Educational and health institutions, health professionals, government health authorities, and policy makers need to be involved in the proper development of policies to improve the information available on the web in order to have a positive impact on the healthy behavior of the population.

Keywords: YouTube; prostate cancer; healthy behavior; exercises; pelvic floor; urinary incontinence

1. Introduction

According to Globocan 2020, prostate cancer is the fourth most common cancer overall, with an incidence of 1,414,259 in men in 2020, exhibiting higher incidence in developed countries.

Prostate cancer is a major cause of disease and mortality among men, with 375,304 men dying of it each year [1,2].

Radical prostatectomy (RP) is a common curative treatment to prevent metastasis. Although mortality after RP is low (5-year survival: 95%), morbidity is high [3]. Surgical treatment involving the removal of the prostate may result in temporary or permanent erectile dysfunction (ED) and urinary incontinence (UI), with considerable impact on quality of life (QoL) [4–7]. While the postoperative incontinence rate is 1% in patients undergoing prostatectomy for benign reasons, a rate between 2% and 66% has been reported after RP [7]. Depending on how continence is defined, almost 80% of men experience incontinence after RP [3]. ED affects 26% to 100% of patients after RP, and the main cause is known to be injury to the neurovascular bundles [5].

Treatment of incontinence involves noninvasive behavioral therapeutic methods consisting of diet modification, bladder training, pelvic floor muscle exercises (PFME), biofeedback, and functional electrical stimulation [4]. Pelvic floor exercises have been used to improve urinary continence following RP, with good results [5,8]. Urinary continence can be achieved through contraction training of the pelvic floor muscles [6]. Pelvic floor muscle training (PFMT) is one of the recommended techniques for the prevention, treatment, and rehabilitation of RP-related complications [5]. It can improve UI and ED after prostatectomy [6]. The aim of PFME, first defined by Arnold Kegel in 1948 as a behavioral therapeutic method for treating incontinence, is to enhance muscle volume and contraction strength in case of increased intra-abdominal pressure [4]. Many studies of PFMT have been conducted in post-prostatectomy patients, delivered both before and after surgery. An increase in speed of recovery was found in more active rehabilitation intervention targeting physiological PFM function of fast- and slow-twitch muscle fibers [9]. However, a major barrier to the success of any training program is adherence, and PFMT is no different. Many patients search the Internet for medical information, but they lack the tools to evaluate the advice provided [10]. As of 3 March 2020, an estimated 58.7% of the global population, representing 4,574,150,134 users, had access to the Internet [11]. YouTube is one of the most popular Internet sites and it also operates as a primary Internet platform for consumer-targeted health information. It is the largest video archive website in the world and attracts 95% of Internet users, with 30 million active users every day [12]. It has 5 billion visits per day and 1 billion hours are watched daily [13]. YouTube and other online video-sharing sites have also become important channels for science popularization and communication [14].

Advances in eHealth technology have cultivated transactional opportunities for patients to access, share, and monitor health information [15]. Digital behavior change interventions use digital health technologies for behavior modification for the maintenance and improvement of health [16], and patients and health professionals tend to search the Internet for information about many health-related topics. In fact, 81% of all Internet users go online and search for information related to health [13]. Generally, patients search for detailed information about recommended exercises. Therefore, the authorship, quality, and validity of the information contained in videos must be considered [12], as some of them present commercial content that may affect the attitude and decision-making of consumers [17].

Video-sharing websites must be understood as social media sites where there is no editorial selection or quality assessment. Likewise, potentially harmful and inaccurate information about science and biomedical topics can be disseminated. The kind of content stored on YouTube and the quality of this information is unclear [18], but a platform such as YouTube has the potential to be an important resource for sharing and disseminating health-related information [19].

One question that has been highlighted by biomedical institutions is whether YouTube provides users and (potential) patients with accurate and helpful information or might the videos possibly be harmful and misleading. Healthcare professionals and organizations should be encouraged to provide more beneficial material and animated videos to people looking for comprehensive, reliable information on the Internet [20]. Healthcare providers and government agencies have expressed their concern about the veracity and quality of the information available on this platform as the health-content videos uploaded by various sources can be misleading and present inaccurate information to patients [20].

2. Materials and Methods

2.1. Search Strategy

On 24 January 2021, a search was conducted on http://www.youtube.com (assessed on 24 January 2021), using the following search terms: 'Prostate cancer–exercises–pelvic floor'.

The first 150 videos available to viewers were selected. The aim was to replicate a simple search strategy that could be conducted by anyone, so the search was not restricted using filters. Hence, YouTube sorted video results by their relevance according to the patented ranking algorithm active on that specific day. All the videos were added to a spreadsheet and then submitted for screening for duplicates, as well as in order to apply the research team's inclusion and exclusion criteria. Exclusion criteria were non-English language, duplicated videos, pelvic floor exercises for women, and/or related to advertisements. Finally, 133 videos were assigned to 2 different examiners who viewed, analyzed, and evaluated them independently over a period of 5 weeks (Figure 1).

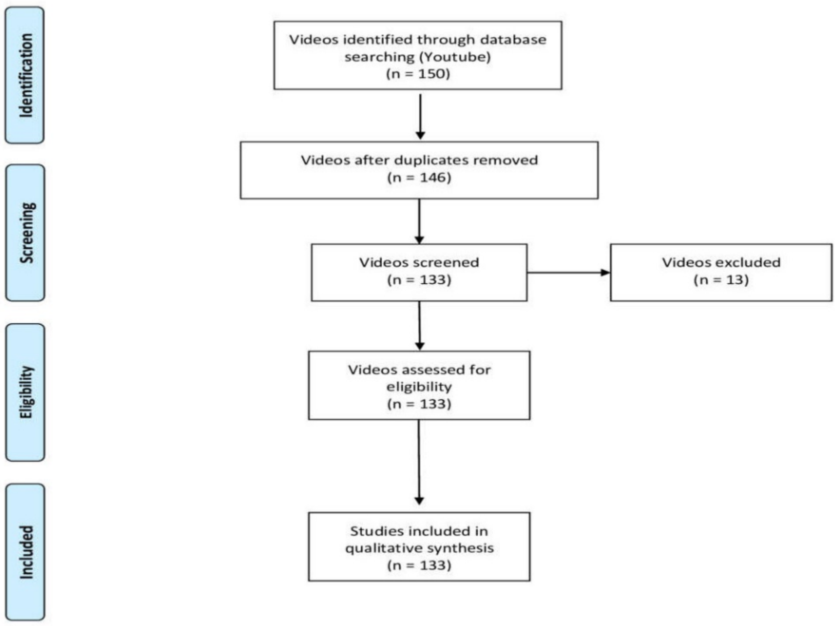


Figure 1. Video flow diagram [21].

2.2. Outcome Measures

Based on their production source, the videos were categorized into six groups: Health Organization (clinic/hospital), Healthcare Professional, Non-Healthcare Professional, Academic Institution, Media (TV, newspaper, etc.), or NGOs. Videos were also coded according to their continent of origin (America, Europe, Asia, Africa, Australia).

Information on when the exercises were recommended (preoperative, postoperative, or both) and the name of the exercises themselves (pelvic floor exercises, Kegel exercises, breathing exercises, yoga exercises, general exercises, or various) was also collected.

The objective of the videos was also identified, based on their indication with regard to UI, fecal incontinence (FI), sexual disfunction (SeD), pelvic pain (PP), and/or incontinence + sexual disfunction (UI + SeD). The quality of the information provided was completed with other collected variables, such as patient position for performing the exercises, number of daily and weekly repetitions, fiber type (phasic or tonic), breathing type, warning about the inappropriate use of accessory muscles, and whether or not it was necessary to use some kind of material to carry out the exercises.

Similarly, descriptive characteristics of each video (View counts, Likes, Dislikes, Origin, Days online, Author, and Duration) were collected. Video popularity was assessed using the Video Power Index (VPI) ((like count/(dislike count + like count)) × 100)) [12,22,23] and View Ratio (Views count/Days online) [24]. The educational quality of the 133 selected videos was determined using the DISCERN tool (Quality Criteria for Consumer Health Information) [25] and the Global Quality Scale (GQS) [26].

A modified 5-point DISCERN tool [27], adapted from the original DISCERN tool for assessment of written health information by Charnock et al. [25], was used for this study. It was created by the Division of Public Health and Primary Care at Oxford University, London, to gauge the quality of information regarding treatment choices for health problems, and was first published in 1999 [25]. The questionnaire consists of a total of 5 questions in addition to an overall quality rating. Each question represents a different quality criterion, rated from 1 to 5 points (1: Very Poor, 2: Poor, 3: Average, 4: High, and 5: Very High quality).

GQS assesses the content quality of online resources. One point is assigned for each of the five identifiable criteria present in a video, with five being the highest educational quality [26]. This scale covers the accessibility and quality of the information, the overall flow of information, and how useful it would be for a user [24].

2.3. Statistical Analysis

Two different analyses were performed on the dataset: The first analysis was based on machine learning techniques, considering a binary classification problem, and thereby dividing the sample into 'Relevant Videos' (C1) or 'Non-Relevant Videos' (C2). These two classes were defined in two cases, using the means of the DISCERN and Principal Component Analysis (PCA) [28]. PCA visualizes the information in a dataset described by multiple interrelated variables. The information in a given dataset corresponds to the total variation. PCA identifies directions (or principal components) along which there is a maximum variance in the data. It is used to extract the important information from a multivariate dataset and express this information as a set of a few new variables, i.e., the principal components. Thus, PCA reduces the dimensionality of multivariate data to two or three principal components (in this case, two principal components were used to group data) that can be visualized graphically, with minimal loss of information.

The objective of the first analysis was to determine the variables that best distinguish these two video classes. The classification problem was studied in two cases, when both classes were defined by the following variables:

1. DISCERN: In this case, samples in the first class, C1 (Relevant Videos), were considered to be the videos whose DISCERN variable value was greater than the mean for this variable, and the rest of the videos were defined as C2.

2. PCA: This was a binary variable that already contained the class of each sample after performing a PCA and a k-means algorithm. In this case, the sample grouping into C1 and C2 is given by the variability in data and there is no 'a priori' information as to the relevance of the videos in each class. Consequently, reducing the number of variables to two (as per PCA) that group the most important information of all the initial variables means that no information is lost.

For each of the two cases, a classification algorithm was applied to attain a minimum list of variables that best assigned the samples into the two groups. Accuracies were determined by Leave-One-Out Cross Validation (LOOCV) using a nearest-neighbor classifier. The discriminatory power of the variables was established according to their Fisher's Ratio (FR). Variables with an elevated FR are highly discriminatory, since they have low intraclass dispersion and high interclass distance. In a binary classification, the FR of the variable *j* is given by:

$$\frac{(\mu_{j1} - \mu_{j2})^2}{{\sigma_{j1}}^2 + {\sigma_{j2}}^2}$$

where μ_{ji} is a measure of the center of mass of the probability distribution of the variable *j* in class *i* (*i* = 1,2), and σ_{ji} is a measure of its dispersion within this class. Variables with a high FR are highly discriminatory, since they have low intraclass dispersion and high interclass distance.

In addition to FR, Pearson, Kendal, and Spearman Correlation Factors of the variables with the defined classes were computed. These factors also reveal the relevance (discriminatory power) of the variables for the classification criteria.

The second analysis was a statistical one using a t-test and Wilcoxon test, illustrating how significant the differences between both classes are from the perspective of each variable. The statistical analysis also shows the relevance of each variable regarding the grouping of the samples into C1 and C2, in each of the two cases. The *t*-test determines whether the means of the variables in both datasets are significantly different from each other. The null hypothesis (H0) assumes that there is no difference between groups. When the statistical analysis yields 1, it means that there is enough evidence to reject this hypothesis, and if it results in 0, the null hypothesis must be accepted. The accepted value level of significance or error probability is alpha < 0.05. Consequently, the Wilcoxon test essentially calculates the difference between each set of pairs and analyzes these differences, whereas H0 refers to the equality of the population medians of two groups of samples (C1 and C2).

3. Results

3.1. Statistical Analysis

Descriptive statistics were obtained from each video, calculating their mean and establishing their minimum and maximum values along with their standard deviation (SD). The same values were obtained for the quality scales.

An ICC analysis was conducted to gauge inter-examiner concordance, with 95% confidence intervals (CI) based on mean rating (k = 2), consistency, two-way random model, and Pearson's Correlation method. Level of significance was set at $p < 0.05$. Inter-rater agreement scored 0.9505 for this study.

3.2. Video Characteristics

Most of the videos (32.3%) were produced by Health Institutions, followed by Healthcare Professional (24.8%), Media sources (18.0%), Non-Healthcare Professional (12.0%), NGOs (6.8%), and Academic Institutions (6.0%) (Figure 2).

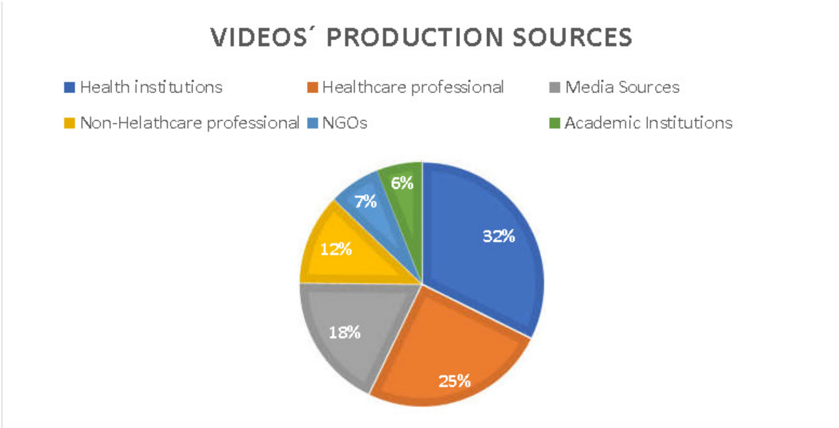


Figure 2. Video production sources.

The origin of the videos was America in 65.4% of cases, Australia (28.6%), Europe (3.8%), Asia (0.8%), and Africa (0.8%).

The mean number of views was 124,354 (SD = 472,419), mean video duration was 14.42 min (SD = 19.34), mean number of days online was 1777 (SD = 1180), mean view ratio was 138.30 (SD = 788.31), mean Likes count was 1082 (SD = 4883), mean Dislikes count was 68.58 (SD = 265.45), and mean VPI was 92.28 (SD = 8.89). DISCERN and GQS scores of both independent observers were averaged to calculate the mean scores of each video. Considering average results, DISCERN scored a mean of 3.35 (SD = 1.25), while GQS scored a mean of 3.38 (SD = 1.02). Table 1 shows the basic descriptive statistics data and DISCERN and GQS results of the videos included.

Table 1. Descriptive statistics.

	Mean	±	SD	Median	Min–Max		
Views	124,354	±	472,419	3678	14	–	3,484,686
Days Online	1777	±	1180	1597	220	–	4518
Views/day	138.30	±	788.31	2.67	0	–	8808
Likes	1082	±	4883	21.00	0	–	48,668
Dislikes	68.58	±	265.45	1.00	0	–	2195
Subscribers	95,039	±	343,699	3570	0	–	2,980,000
VPI	92.28	±	8.89	94.38	50.00	–	100.00
DISCERN 1	0.98	±	0.15	1.00	0	–	1
DISCERN 2	0.59	±	0.49	1.00	0	–	1
DISCERN 3	0.89	±	0.31	1.00	0	–	1
DISCERN 4	0.53	±	0.50	1.00	0	–	1
DISCERN 5	0.36	±	0.48	0.00	0	–	1
DISCERN Total	3.35	±	1.25	3.00	0	–	5
GQS	3.38	±	1.02	4.00	1	–	5

According to the mean DISCERN scores of both observers, the quality of the videos was found to be Very Poor in 5.3% of the cases, Poor in 24.1%, Average in 21.1%, High in 27.8%, and Very High in 21.8%. However, referring to the mean GQS, the quality of the videos was found to be Very Poor in 3.8% of the cases, Poor in 17.3%, Average in 27.1%, High in 40.6%, and Very High in 11.3%.

As for the relationship between the production source and educational quality of the videos, according to DISCERN, the videos made by NGOs attained higher scores (4.33), followed by Health Institutions (3.88) and Academic Institutions (3.88), Healthcare Professional (3.45), Media (2.79), and finally those created by a Non-Healthcare Professional (1.75). Likewise, the GQS (the other variable that alludes to educational quality) ranked the origins in the same order: NGO's and Academic Institutions also attained higher scores (4.00), followed by Health Institutions (3.77), Healthcare Professional (3.45), Media (2.96), and, finally, Non-Healthcare Professional (2.19). These data and the views ratio can be seen in Table 2.

Table 2. DISCERN and GQS scores.

	DISCERN Score	GQS Score
Academic Institution	3.88	4.00
Media	2.79	2.96
NGO	4.33	4.00
Health Institutions	3.88	3.77
Non-Healthcare Professional	1.75	2.19
Healthcare Professional	3.45	3.45
Total	3.35	3.38
Africa	0.00	1.00
America	3.33	3.36
Asia	3.00	3.00
Australia	3.55	3.55
Europe	3.40	3.40
Total	3.35	3.38

The same table shows that higher scores on the DISCERN scale were assigned to videos from Australia (3.55), followed by Europe (3.40), America (3.33), Asia (3.00), and Africa (0.00). Meanwhile, the GQS scale assigned almost the same scores in videos made in Australia (3.55), followed by Europe (3.40), America (3.36), Asia (3.00), and, finally, Africa (1.00).

VPI mean scores for each of the five categories in which the DISCERN and GQS variables classified the videos were also analyzed. Videos with a Very Poor DISCERN score had a VPI mean of 89.20, videos classified as Poor had a VPI of 92.75, Average videos had 92.11, High videos had 89.13, and videos classified as Very High for DISCERN had 94.83. Likewise, videos with Very Poor GQS score had a VPI mean of 89.88, videos classified as Poor had a VPI of 90.94, Average videos had 93.31, High videos had 90.30, and videos classified as Very High for GQS had 96.41.

Regarding the objective of the videos, 39.85% of them did not specify their objective, 26.32% aimed to treat UI + SeD, 23.31% UI, 5.26% SeD, 4.51% PP, and 0.75% FI.

The terms used by the authors to denominate the exercises in the videos were Pelvic Floor Exercises (PFE) in 42.11% of cases, followed by Kegel Exercises (25.56%), Undefined (17.29%), and General Exercises (15.04%).

To carry out the exercises, most authors did not indicate the need for specific complimentary materials (65.4%), 24.1% suggested some type of domestic material, and only 10.5% of videos indicated the use of specific material, such as balls, special benches, or gym equipment.

Regarding the period of time over which the authors recommended doing the exercises, most (48.1%) did not specify it, followed by 30.1% who recommend the exercises for the postoperative period, 18.0% for the pre- and post-operative period, and, finally, 3.8%, for the preoperative one.

Of the videos, 78.9% did not specify the frequency of performing the exercises, while 16.6% suggested a frequency of three times a day, 3.8% indicated once a day, and 0.8% recommended two times a day.

The position of the patient during the performance of the exercises was not specified in 72.9% of the videos, in 15% various positions were indicated, while only 6.0% of videos recommended doing them sitting down, supine in 5.3%, and standing up in 0.8% of cases. Regarding breathing during exercises, 91.7% of videos did not mention it, 5.3% referred to breathing but did not indicate how to do it, while 3.0% explained that it was to be done during exhalation. Finally, 94% of videos did not indicate anything about the use of accessory muscles, whereas 6% of them recommended not using them. Regarding the effort on different types of muscle fibers, 21.8% of videos recommended the work on the tonic fiber and 9.77% the phasic fiber. The rest of the videos made no recommendation concerning this topic.

4. Discussion

This study was undertaken to better comprehend the nature of evidence independently accessed by patients on a massive online, media-sharing platform. The increasing use of new technologies and vast amount of content available have turned the Internet into an important source of health information, for healthcare professionals (who use new technologies more frequently to communicate with patients and achieve the necessary adherence and interaction in health processes) as well as patients (who use it as a source of information). However, this entails some risk, in as much as the information available on the web is not subject to any kind of inspection. Consequently, this can lead to erroneous and even harmful messages being conveyed to users and/or viewers. In the search process to improve health processes, rigorous education and dissemination are necessary, preferably carried out by health professionals. In the case of the videos analyzed, most (32.3%) were produced by Health Institutions, which can be considered as a guarantee, due to the fact that they are in the second place of the quality scales.

Very few videos were produced by NGOs (6.8%), even though the quality of their productions obtained the highest scores, which is in agreement with Masefield et al., who observed that NGO data constitute a vast and valuable source of information for health policy makers and public health systems research [29]. Videos whose source of production is Non-Health Professionals are those that obtained the worst scores on these scales.

Unfortunately, the percentage of videos made by Academic Institutions was the lowest (6.0%), a result that is in line with other studies that claim that Health Institutions are underrepresented in the publication of videos of medical information [22]. Taking into account the fact that the quality of the videos that they produce is one of the highest, they should consider academic projects to increase their presence and flood the Internet with high-quality information. As in similar studies, it was difficult in many cases to find out who had developed the video content, as well as for viewers to assess the reputation of the sources [30].

The average length of the videos in our study was 14.42 min, which is higher than previous studies that reported an average duration of 6.17–10.35 min [22,31]. No statistically significant relationship was found between video length and their quality or popularity, contrary to the study by Akif Aydin that found that the videos with the longest duration and the highest VPI appeared to be associated with higher quality scores [23].

ICC is a widely used test–retest, intra-rater, and inter-rater reliability index. Based on the 95% confident interval of the ICC estimation [32], the inter-reviewer agreement for this study was 0.9505, pointing toward Excellent concordance between both examiners, according to the ICC intervals shown in Table 3.

Table 3. ICC intervals.

Interval	Reliability *
Less than 0.5	Poor
0.5–0.75	Moderate
0.75–0.90	Good
Greater than 0.90	Excellent

* Based on the 95% confidence interval.

PC incidence rates are highly variable worldwide. Research has shown that African American men have the highest incidence of prostate cancer worldwide [33], while Chu et al. [34] reported that incidence rates of prostate cancer were as much as 40 times higher among African American men than those in Africa. In 2018, the highest mortality rates were recorded in Central America (10.7 per 100,000 people), followed by Australia and New Zealand (10.2), and Western Europe (10.1) [33]. As found in this study, most videos were produced in America (65.4%), where, in 2021, the estimated number of new cases was 248,530 [2], followed by those produced in Australia (28.6%), which points towards countries with higher incidence and mortality as those that produce more videos about PC. The overall means of DISCERN and GQS indicate that the YouTube contents analyzed have medium–high reliability and quality. They also show higher scores for videos produced in Australia, Europe, and America.

A statistically significant relationship between DISCERN and GQS scales was also detected (Pearson Coefficient = 0.9023; $p < 0.001$). Analyzing the scores obtained to appraise the educational quality of videos with both scales, it should be noted that most videos were of High quality (27.8% and 40.6% of cases, for DISCERN and GQS scales, respectively), Very High in 21.8% (DISCERN) and 11.3% (GQS), Average quality was found in 21.1% (DISCERN) and 27.1% (GQS), Poor quality in 24.1% (DISCERN) and 17.3% (GQS), and Very Poor educational quality in 5.3% (DISCERN) and 3.8% (GQS) of cases. This contrasts with previous studies that identified a worrying amount of poor-quality and erroneous information in YouTube videos related to health [17,23,24]. The results shown in this study indicate, as other authors have already mentioned, that YouTube is a solid and useful source of information [35], and different ways are needed to ensure that users access the most appropriate ones.

Since there is no mark to identify high-quality videos, it is recommended for their easy identification that patients attend to the source of production. Since Academic Institutions, NGOs, and/or Health Institutions show higher figures in quality scales, these should be the reference. It would also be advisable to discard those videos that contain commercials and could present any conflict of interest.

From the perspective of the PCA, if the coordinates of each video are graphically represented in a coordinate system in which the axes refer to the principal components, the videos are naturally grouped into two groups of points, without any manual intervention or possible bias. These two groups, which can be seen in Figure 3, collect the videos with higher ratings in DISCERN and GQS in one group (C1) and those with lower ratings in the other group (C2) (Figure 3).

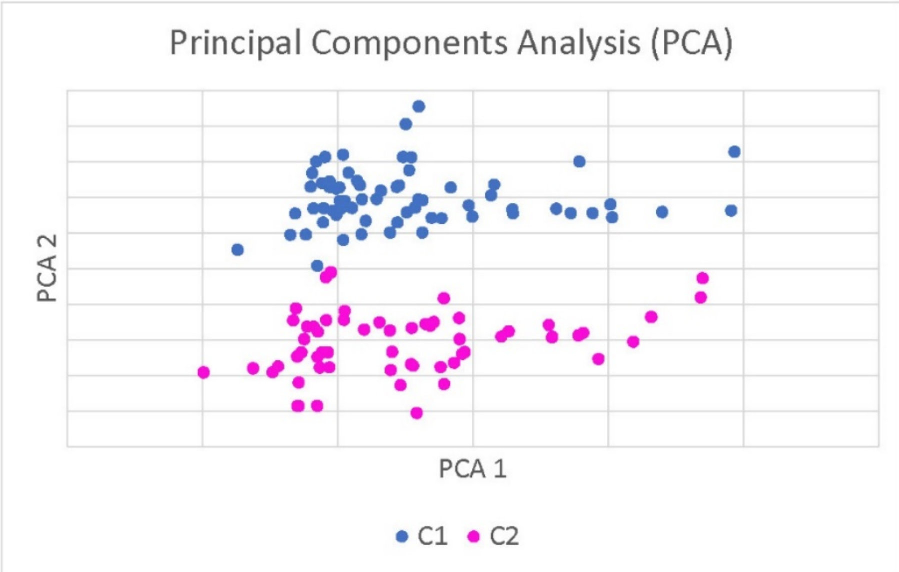


Figure 3. Extraction method: Principal Axis Factorization. Rotation method: Varimax Rotation with Kaiser Normalization.

The statistical power of these two combined analyses indicates that the PCA enables the videos to classify themselves in two groups (C1, high-quality videos, and C2, low-quality videos) without any information being lost on any of the variables considered in the entire study. Consequently, the quality of the videos can be analyzed across the two principal components, rather than having to do so across all the variables in the study because the PCAs have a high statistical relationship with the variables evaluated. This is a huge advantage when carrying out this kind of analysis, considering that no important information is lost through these two principal components because their level of relationship with the main variables that evaluate the quality of the videos is maximum, and *p*-values reveal a high, statistically significant relationship.

The statistical power of these combined analyses is such that it enables the videos' classification, according to their quality, by considering two variables (principal components) that combine together the information from all the other variables analyzed in the study.

Health Institutions and Healthcare Professionals refer more rigorously to the correct terminology for these types of exercises than the rest of the sources, who refer to them as exercises (without specifying the term). Using appropriate terminology, so that patients know the conceptual distinctions, is necessary due to the increasing culturalization of society. People are willing to feel as if they are "decision-makers" and not only "patients", so they demand a quota of self-resolution of their needs, with its advantages and disadvantages, such as self-taught training through the new online resources. To carry out this learning, it is necessary to use appropriate terminology for each type of situation [36] and thus participate in the development of the patient's own decision-making capacities, with explanatory models that enable "experience" of the disease to be acquired. The way of narrating the disease is not only an academic finding, since it has practical significance in education, health promotion, and the development of institutional and professional devices [37].

Most of the videos did not indicate the need for extra material to carry out the exercises, which facilitates their implementation at home.

Regarding the indication of when to perform the exercises, most authors did not specify it, followed by a high percentage who recommended performing them post-surgery. The lowest percentage (3.8%) corresponds to those that indicated its pre-operative performance. The data obtained do not agree with studies such as that of Milios et al., who showed that a PFME program started prior to prostate surgery enhanced post-surgical measures of pelvic floor muscle function, reduced post-prostatectomy incontinence, and improved QoL outcomes related to incontinence [9].

Evidence shows that daily performance frequency is important, with two sets of PFM exercises per day as the most highly recommended, in accordance with the Glazener study, which recommends pelvic floor contractions twice per day [38], or the study of Kraemer and Ratamess, who reported greater improvements when participants exercised twice a day compared to only once per day [39]. After analyzing the videos, it was observed that 78.9% of the videos did not mention the frequency, and only 16.5% indicated that they were to be performed 3 times a day, as other authors also indicate [9].

The most appropriate exercises should be performed in three positions, sitting down, standing up, and supine positions, during all sessions [4,38]. However, most videos did not mention the position, while only 15% indicated the need to perform them in several positions.

Breathing during exercises is very important, and subjects should not hold their breath during pelvic floor activation, avoiding the Valsalva maneuver [4]. Despite this fact, in most of the videos, it was not mentioned (91.7%), or if it was, no indication was made of how to do it (5.3%). Only a minimum (3%) correctly explained how to do contractions during prolonged expiration [8].

Most authors indicate that activation of the superficial abdominal muscles should be reduced by means of visual and tactile pointers, and accessory muscles should not be used when performing exercises, as indicated by Gómez Lanza [40]. Despite this, 94% of videos did not include any information on this topic.

The indication on fast- and slow-twitch training was not represented in the videos either, since only 21.8% of videos recommended work on slow twitch and 9.77% did so on fast twitch, whereas it is necessary to work on each of these types [9].

Limitations

Adhering to the principles of simulating what users see, no searches were conducted under incognito mode so as to avoid the influence of browsing history and geographical locations.

Although YouTube content changes over time, our analysis represents the state of the videos at one specific time [41].

Only videos directly accessed on YouTube upon searching for 'Prostate cancer–exercises–pelvic floor' were included in this study. External links from other medical-related websites were excluded in our analyses.

As a massive Internet-based platform, YouTube searches are continuously evolving as new videos are constantly being uploaded, viewed, and rated. Furthermore, our search was limited to the first 150 videos. As other studies have previously explained, our methodology was designed to replicate the average patient's search attempt, since most Internet users do not look past the first 50 search results [35].

5. Conclusions

One of the main barriers to the success of any training program is adherence, and different methods are needed to ensure that users access the most appropriate ones. YouTube can be a strategy to achieve this. It is a solid and useful source of information, but a process of verification and validation of the information available on the web is necessary, as well as educational programs to enable society to access the most reliable information. YouTube is one of the most important development tools of the eHealth era because videos are a simple, attractive, and affordable source of information, with great dissemination capacity available to the population. The quality of the videos available on YouTube regarding the recommended exercises for pelvic floor in PC surgery, according to the scores obtained, is High.

Educational and Health Institutions, as well as Health Professionals, Government Health Authorities, and legislators, must be involved in the correct development of policies that improve the information available on the web in order to generate a positive impact on the healthy behavior of the population.

Future research should develop tools to make it easier for patients to identify the highest quality videos from a health perspective.

Author Contributions: Conceptualization, A.M.R.-R. and M.B.-D.; methodology, J.C.; software, P.L.-D. and M.d.l.F.-C.; validation, I.E.-P. and M.B.-D.; formal analysis, A.M.R.-R.; investigation, J.C.; resources, M.C.S.-F.; data curation, P.L.D. and M.d.l.F.-C.; original writing—A.M.R.-R. and M.B.-D.; review and editing—J.C.; visualization, I.E.-P. and M.C.S.-F.; supervision, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. The Global Cancer Observatory. All Cancers. 2020. Available online: https://gco.iarc.fr/today (accessed on 23 June 2021).
2. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2021. *CA Cancer J. Clin.* **2021**, *71*, 7–33, doi:10.3322/caac.21654.
3. Hodges, P.W.; Stafford, R.E.; Hall, L.; Neumann, P.; Morrison, S.; Frawley, H.; Doorbar-Baptist, S.; Nahon, I.; Crow, J.; Thompson, J.; et al. Reconsideration of pelvic floor muscle training to prevent and treat incontinence after radical prostatectomy. *Urol. Oncol. Semin. Orig. Investig.* **2020**, *38*, 354–371, doi:10.1016/J.UROLONC.2019.12.007.
4. Aydın Sayılan, A.; Özbaş, A. The effect of pelvic floor muscle training on incontinence problems after radical prostatectomy. *Am. J. Mens Health.* **2018**, *12*, 1007–1015, doi:10.1177/1557988318757242.
5. De Lira, G.H.S.; Fornari, A.; Cardoso, L.F.; Aranchipe, M.; Kretiska, C.; Rhoden, E.L. Effects of perioperative pelvic floor muscle training on early recovery of urinary continence and erectile function in men undergoing radical prostatectomy: A randomized clinical trial. *Int. Braz. J. Urol.* **2019**, *45*, 1196–1203, doi:10.1590/S1677-5538.IBJU.2019.0238.
6. Pan, L.H.; Lin, M.H.; Pang, S.T.; Wang, J.; Shih, W.M. Improvement of urinary incontinence, life impact, and depression and anxiety with modified pelvic floor muscle training after radical prostatectomy. *Am. J. Mens Health.* **2019**, *13*, doi:10.1177/1557988319851618.
7. O'Callaghan, M.E.; Raymond, E.; Campbell, J.; Vincent, A.D.; Beckmann, K.; Roder, D.; Evans, S.; McNeil, J.; Millar, J.; Zalcberg, J.; et al. Tools for predicting patient-reported outcomes in prostate cancer patients undergoing radical prostatectomy: A systematic review of prognostic accuracy and validity. *Prostate Cancer Prostatic Dis.* **2017**, *20*, 378–388, doi:10.1038/pcan.2017.28.
8. Prota, C.; Gomes, C.M.; Ribeiro, L.H.S.; De Bessa, J.J.; Nakano, E.; Dall'Oglio, M.F.; Bruschini, H.; Srougi, M. Early postoperative pelvic-floor biofeedback improves erectile function in men undergoing radical prostatectomy: A prospective, randomized, controlled trial. *Int. J. Impot. Res.* **2012**, *24*, 174–178, doi:10.1038/ijir.2012.11.
9. Milios JE, Ackland TR, Green DJ. Pelvic floor muscle training in radical prostatectomy: a randomized controlled trial of the impacts on pelvic floor muscle function and urinary incontinence. *BMC Urol.* **2019**, *19*, 1–10, doi:10.1186/s12894-019-0546-5.
10. Drozd, B.; Couvillon, E.; Suarez, A. Medical YouTube videos and methods of evaluation: Literature review. *J. Med. Internet Res.* **2018**, *4*, e8527, doi:10.2196/mededu.8527.
11. World Internet Users Statistics and 2020 World Population Stats. 2020. Available online: https://www.internetworldstats.com/stats.htm (accessed on 13 April 2020).
12. Atci, A.G. Quality and reliability of the information on YouTube videos about Botox injection on spasticity. *Rom. Neurosurg.* **2019**, *33*, 473–477, doi:10.33962/roneuro-2019-075.
13. Omnicore Agency. YouTube by the Numbers: Stats, Demographics & Fun Facts. 2020. Available online: https://www.omnicoreagency.com/youtube-statistics/ (accessed on 13 April 2020).
14. Allgaier, J. Science and medicine on YouTube. In *Second International Handbook of Internet Research*; Springer: Amsterdam, The Netherlands, 2020; pp. 7–27, doi:10.1007/978-94-024-1555-1_1.
15. Paige, S.R.; Stellefson, M.; Krieger, J.L.; Anderson-Lewis, C.; Cheong, J.; Stopka, C. Proposing a transactional model of eHealth literacy: Concept analysis. *J. Med. Internet Res.* **2018**, *20*, e10175, doi:10.2196/10175.

16. Craig, K.J.T.; Morgan, L.C.; Chen, C.-H.; Michie, S.; Fusco, N.; Snowdon, J.L.; Scheufele, E.; Gagliardi, T.; Sill, S. Systematic review of context-aware digital behavior change interventions to improve health. *Transl. Behav. Med.* **2021**, *11*, 1037–1048, doi:10.1093/tbm/ibaa099.
17. Loeb, S.; Sengupta, S.; Butaney, M.; Macaluso, J.N.; Czarniecki, S.W.; Robbins, R.; Braithwaite, R.S.; Gao, L.; Byrne, N.; Walter, D.; et al. Dissemination of misinformative and biased information about prostate cancer on YouTube. *Eur. Urol.* **2019**, *75*, 564–567, doi:10.1016/j.eururo.2018.10.056.
18. Hartley, J. *Digital Futures for Cultural and Media Studies*; Wiley: Hoboken, NJ, USA, 2012; doi:10.1002/9781118106723.
19. Madathil, K.C.; Rivera-Rodriguez, A.J.; Greenstein, J.S.; Gramopadhye, A.K. Healthcare information on YouTube: A systematic review. *Health Inform. J.* **2015**, *21*, 173–194, doi:10.1177/1460458213512220.
20. Aydin, M.F.; Aydin, M.A. Quality and reliability of information available on YouTube and Google pertaining gastroesophageal reflux disease. *Int. J. Med. Inform.* **2020**, *137*, 104107, doi:10.1016/J.IJMEDINF.2020.104107.
21. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med.* **2009**, *6*, e1000097, doi:10.1371/journal.pmed1000097.
22. Kuru, T.; Erken, H.Y. Evaluation of the quality and reliability of YouTube videos on rotator cuff tears. *Cureus* **2020**, *12*, e6852, doi:10.7759/cureus.6852.
23. Aydin, M.A.; Akyol, H. Quality of information available on YouTube videos pertaining to thyroid cancer. *J. Cancer Educ.* **2019**, *18*, 1–14, doi:10.1007/s13187-019-01502-9.
24. Erdem, M.N.; Karaca, S. Evaluating the accuracy and quality of the information in kyphosis videos shared on YouTube. *Spine* **2018**, *43*, E1334–E1339, doi:10.1097/BRS.0000000000002691.
25. Charnock, D.; Shepperd, S.; Needham, G.; Gann, R. DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *J. Epidemiol. Community Heal.* **1999**, *53*, 105–111, doi:10.1136/jech.53.2.105.
26. Kunze, K.N.; Cohn, M.R.; Wakefield, C.; Hamati, F.; LaPrade, R.F.; Forsythe, B.; Yanke, A.B.; Chahla, J. YouTube as a source of information about the posterior cruciate ligament: A content-quality and reliability analysis. *Arthrosc. Sport Med. Rehabil.* **2019**, *1*, e109–e114, doi:10.1016/j.asmr.2019.09.003.
27. Esen, E.; Aslan, M.; Sonbahar, B.Ç.; Kerimoğlu, R.S. YouTube English videos as a source of information on breast self-examination. *Breast Cancer Res. Treat.* **2019**, *173*, 629–635, doi:10.1007/s10549-018-5044-z.
28. Zhang, Z.; Castelló, A. Principal components analysis in clinical studies. *Ann. Transl. Med.* **2017**, *5*, 351–351, doi:10.21037/atm.2017.07.12.
29. Masefield, S.C.; Megaw, A.; Barlow, M.; White, P.C.L.; Altink, H.; Grugel, J. Repurposing NGO data for better research outcomes: A scoping review of the use and secondary analysis of NGO data in health policy and systems research. *Heal. Res. Policy Syst.* **2020**, *18*, 1–22, doi:10.1186/s12961-020-00577-x.
30. Heathcote, L.C.; Pate, J.W.; Leake, H.B.; Moseley, L.; Kronman, C.A.; Fischer, M.; Timmers, I.; Simons, L.E. Pain neuroscience education on YouTube. *PeerJ* **2019**, *7*, e6603, doi:10.7717/peerj.6603.
31. Gokcen, H.B.; Gumussuyu, G. A quality analysis of disc herniation videos on YouTube. *World Neurosurg.* **2019**, *124*, 799–804, doi:10.1016/j.wneu.2019.01.146.
32. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163, doi:10.1016/j.jcm.2016.02.012.
33. Rawla, P. Epidemiology of prostate cancer. *World Oncol.* **2019**, *2*, 63–89, doi:10.14740/wjon1191.
34. Chu, L.W.; Ritchey, J.; Devesa, S.S.; Quraishi, S.M.; Zhang, H.; Hsing, A.W. Prostate cancer incidence rates in Africa. *Prostate Cancer* **2011**, *2011*, 1–6, doi:10.1155/2011/947870.
35. Steinberg, P.L.; Wason, S.; Stern, J.M.; Deters, L.; Kowal, B.; Seigne, J. YouTube as source of prostate cancer information. *Urology* **2010**, *75*, 619–622, doi:10.1016/j.urology.2008.07.059.
36. Boixareu, R.M. *De la Antropología Filosófica a la Antropología de la Salud*; Herder Editorial: Barcelona, Spain, 2008.
37. Valverde Jiménez, M.R. Mediación intercultural en el ámbito de la salud. *Enferm. Glob.* **2013**, *12*, 383–390.
38. Glazener, C.; Boachie, C.; Buckley, B.; Cochran, C.; Dorey, G.; Grant, A.; Hagen, S.; Kilonzo, M.; McDonald, A.; McPherson, G.; et al. Conservative treatment for urinary incontinence in Men After Prostate Surgery (MAPS): Two parallel randomised controlled trials. *Health Technol. Assess.* **2011**, *15*, 1–296, doi:10.3310/hta15240.
39. Kraemer, W.J.; Ratamess, N.A. Fundamentals of resistance training: Progression and exercise prescription. *Med. Sci. Sports Exerc.* **2004**, *36*, 674–688, doi:10.1249/01.MSS.0000121945.36635.61.
40. Gómez Lanza, E.; Granda Contijoch, M.; Batista Miranda, J.E. Ejercicios Del Suelo Pélvico. *Arch. Esp. Urol.* **2009**, *62*, 889–895.
41. Radonjic, A.; Fat Hing, N.N.; Harlock, J.; Naji, F. YouTube as a source of patient information for abdominal aortic aneurysms. *J. Vasc. Surg.* **2020**, *71*, 637–644, doi:10.1016/j.jvs.2019.08.230.

8.3. Article 3

“Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown”

Citation:

Rodriguez-Rodriguez, A.M.; Blanco-Diaz, M.; de la Fuente-Costa, M.; Hernandez-Sanchez, S.; Escobio-Prieto, I.; Casaña, J. Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown. Int. J. Environ. Res. Public Health 2022, 19, 8016. https://doi.org/10.3390/ijerph19138016

International Journal of *Environmental Research and Public Health*

MDPI

Review

Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown

Alvaro Manuel Rodriguez-Rodriguez [1], Maria Blanco-Diaz [2,*], Marta de la Fuente-Costa [2], Sergio Hernandez-Sanchez [3], Isabel Escobio-Prieto [4] and Jose Casaña [1]

[1] Exercise Intervention for Health Research Group (EXINH-RG), Department of Physiotherapy, University of Valencia, 46010 Valencia, Spain; alvaro.manuel.rodriguez@gmail.com (A.M.R.-R.); josecasana@uv.es (J.C.)

[2] Physiotherapy and Translational Research Group (FINTRA-RG), Institute of Health Research of the Principality of Asturias (ISPA), Faculty of Medicine and Health Sciences, University of Oviedo, 33006 Oviedo, Spain; fuentemarta@uniovi.es

[3] Center for Translational Research in Physiotherapy, Department of Pathology and Surgery, Physiotherapy Area, Miguel Hernandez University, 03550 San Juan de Alicante, Spain; sehesa@umh.es

[4] Departamento de Fisioterapia, Facultad de Enfermería, Fisioterapia y Podología, Universidad de Sevilla, 41009 Sevilla, Spain; iescobio@us.es

* Correspondence: blancomaria@uniovi.es; Tel.: +34-667-681-858

Abstract: Background: The world is experiencing a pandemic caused by COVID-19. Insufficient physical activity can increase the risk of illness. Trying to replicate a normal search that any user/patient could do in YouTube, the objective of this study was to evaluate the quality of YouTube videos related to home exercises during lockdown and their adherence to World Health Organization (WHO) recommendations. Methods: A simple search was carried out on YouTube. The first 150 videos were selected. After applying exclusion criteria, 68 videos were analyzed and evaluated. Two statistical analyses based on machine learning techniques were carried out. Videos were classified according to principal component analysis (PCA) models as 'Relevant' and 'Non-Relevant'. Popularity was assessed using the video power index (VPI). Information's quality and accuracy were gauged using the DISCERN scale and global quality score (GQS). Reliability and credibility of information that can be found on health-related websites was assessed using the Health On the Net Code (HONCode). Exercises were evaluated according to WHO recommendations. Results: DISCERN, HONCode, and GQS scored a mean of 2.29, 58.95, and 2.32, respectively. The PCA calculation allowed videos to auto-classify into high- and low-quality videos. Conclusions: The quality of YouTube videos recommending exercises during lockdown is low and doesn't reflect WHO recommendations. Effective strategies and tools capable of indicating the quality of this information are needed to filter out erroneous or non-rigorous information that may affect people's health. These tools should help any user/viewer to distinguish videos of high and low quality.

Keywords: sedentary behaviors; exercises; World Health Organization; COVID-19; YouTube; physical activity; lockdown; health promotion

check for updates

Citation: Rodriguez-Rodriguez, A.M.; Blanco-Diaz, M.; de la Fuente-Costa, M.; Hernandez-Sanchez, S.; Escobio-Prieto, I.; Casaña, J. Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown. *Int. J. Environ. Res. Public Health* **2022**, *19*, 8016. https://doi.org/10.3390/ijerph19138016

Academic Editor: Ellen Glickman

Received: 13 May 2022
Accepted: 28 June 2022
Published: 30 June 2022

1. Introduction

By the end of 2019, a novel coronavirus known as SARS-CoV-2 (COVID-19) suddenly arose in Wuhan, China [1]. This virus manifests as pneumonia due to the fact that it attacks the lower part of the respiratory tract in humans [2]. An international public health emergency was declared on 31 January 2020. As of 14 April 2022, COVID-19 caused over 500,186,525 confirmed cases and over 6,190,349 deaths after being spread worldwide [3].

Most countries have adopted mandatory home lockdown policies. However, prolonged periods of time at home can make staying physically active a major challenge [4]. WHO defines physical activity (PA) as any movement of the body produced by its muscles that requires energy consumption, including exercise and other activities that include physical movement and are conducted as part of play, work, active transport, household chores, and leisure activities [5]. Self-quarantine and prolonged stays at home could be sources of added stress and could also pose challenges for citizens' mental health, contributing to anxiety and depression symptoms [6], as well as increase other health risk behaviors. During the COVID-19 [7] lockdown, a proactive health strategy should be focused on avoiding sedentary behavior. In addition, insufficient PA constitutes the fourth most important risk factor for mortality (6% of deaths worldwide) and severely influences the prevalence of non-communicable diseases (for example, cardiovascular diseases, diabetes, or cancer). It also has other risk factors, such as hypertension, blood glucose excess, or becoming overweight [8]. In fact, insufficiently active people have an increased risk of death of 20% to 30%, compared to people who are active enough and have worsened wellbeing and quality of life [4].

Very few guidelines targeting public health in general consider PA routines on a daily basis for people that live in different levels of isolation during the COVID-19 pandemic [9]. Home exercise, through simple, safe, and easily implementable exercises, is well recommended to maintain good fitness levels. These exercises should include, but not be limited to, activities for strengthening, stretching, and improving balance and control and/or a combination of them [10]. WHO recommends, in a guide published on its website, for confined people without any respiratory illness or suspect of it, 150 min a week of moderated-intensity PA or 75 min a week of higher-intensity PA, or both options combined. In addition, WHO recommends to "*Follow an online exercise class. Take advantage of the wealth of online exercise classes. Many of these are free and can be found on YouTube. If you have no experience performing these exercises, be cautious and aware of your own limitations*" [4].

The use of exercise videos and eHealth, oriented to the encouragement and the delivery of PA through internet, social networks, TV, and mobile technologies are feasible options to maintain mental health and physical fitness during these important periods [10]. As of 3 March 2020, an estimated 58.7% of the world population had access to the internet [11]. In this situation, YouTube attracts 95.0% of all internet users. With much more than 30 million daily users, it is the biggest video website in the world [12]. It has 122 million daily active users, 1 billion hours watched daily, and it attracts about 44% of all internet users [13]. However, there is no data about the quality of the available eHealth and exercise videos, which is particularly relevant in this lockdown period due to COVID-19. Uploaded videos with health content coming from different sources present the risk of showing misleading and inaccurate information to users [14], and the authorship, the quality, and the validity of the information in the videos are essential topics to be considered [12]. In fact, many healthcare professionals and different public authorities have expressed their concerns about the quality and the veracity of the information that can be found on this website.

The objective of this study was to evaluate the quality of YouTube videos related with home exercises during lockdown and their adherence to World Health Organization (WHO) recommendations.

2. Materials and Methods

2.1. Search Strategy

On 3 November 2020 a search on https://www.youtube.com (accessed on 3 November 2020) was carried out using this specific search term: "*COVID Exercises Home*". Our aim was to evaluate the quality of YouTube videos related to home exercises during lockdown and their adherence to WHO recommendations after replicating a simple search process that could be performed by any individual user. No filters were applied in order to avoid limiting the search, so YouTube was permitted to sort video results by relevance in accordance to its ranking algorithm active on that particular day.

The 150 videos that appeared in first order for viewers were considered for this study. They were listed in a spreadsheet for coding them (using the video URLs) and submitted to a duplicates screening (21 duplicated videos were excluded). The exclusion criteria were also applied by the researchers.

The exclusion criteria were applied to the 129 videos that remained after duplicates were removed. These exclusion criteria were: (a) non-English language videos were removed; (b) videos that didn't show exercises were not considered; (c) videos with advertisements were also removed. During the analysis, one of the videos was removed by the platform, so the final result was 61 excluded videos. Finally, 68 videos were independently viewed, analyzed and evaluated by two different researchers (Figure 1). These examiners were each members of a research group at their corresponding universities with long and intensive experience in health topics research.

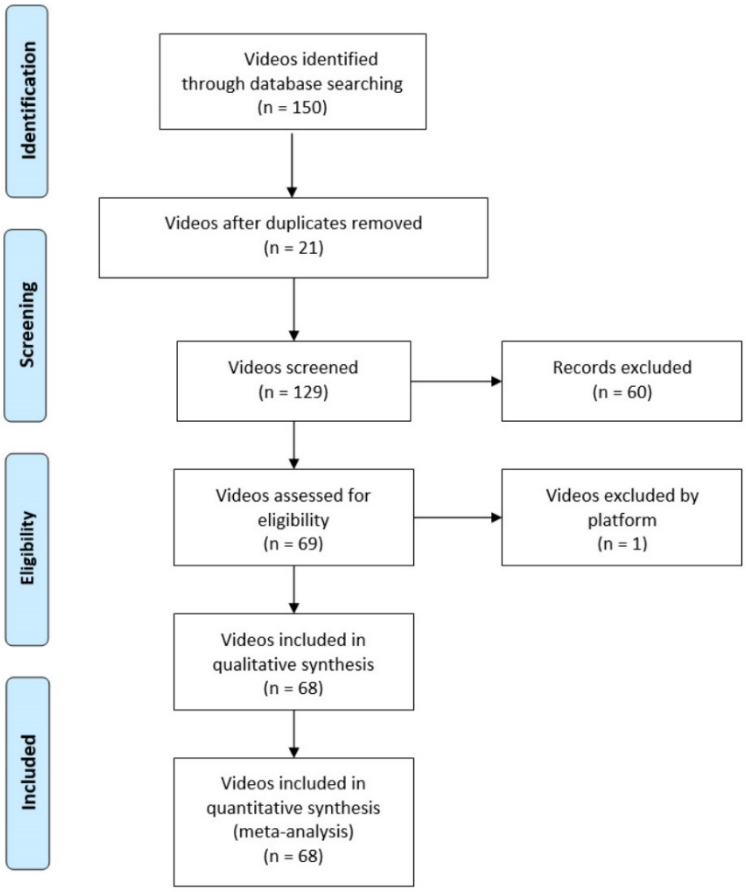


Figure 1. Flow diagram. From [15].

2.2. Outcome Measures

For each video, its descriptive characteristics were collected, such as number of views (view counts), likes and dislikes received, continent of origin, number of days online and their publication date, the author of the video, and its duration (length in minutes and seconds).

Based on their author (source of production), videos were classified into seven groups: health organizations (hospital/clinic), healthcare workers, non-healthcare workers, academic institutions, media (newspaper, TV...), non-governmental organizations (NGO), and sports institutions.

The data referring to the type of target audiences were collected as children, adults, elderly (third age), or all of them, and it was noted if the user had any previous pathology.

Videos were also codified according to their continent of origin (Australia, America, Africa, Asia, and Europe), if the exercises required any kind of material (help/support) to carry them out (professional or home), and the type of exercise (physical, psychic, or physical and psychic).

Exercise time was calculated by multiplying days per week and the recommended exercise time on each video and was classified according to meeting WHO recommendations for exercises during lockdown (150 min/week of moderate exercise, 75 min/week of vigorous exercise or both combined) [4]. The number of exercises contained in each video that agreed with the 12 exercises during lockdown recommended by WHO were also collected [4].

Video popularity was determined by the use of the video power index (VPI) (likes count/(dislikes count + likes count)) $\times$ 100)] [12,14,16] and the view ratio (VR) (view counts per days online) [17].

The reliability and the educative quality of the 68 selected videos were evaluated by the DISCERN [18] scale and the global quality scale (GQS) [19]. A modified five point DISCERN tool [20] was used, which was adapted from the original. It consists of 5 different questions, and one point will be received if the video fulfills that topic or zero points if it does not. The original questionnaire known as the "Quality Criteria for Consumer Health Information" was developed by the "Public Health and Primary Care Division" of Oxford University (London) to assess the information quality of treatment choices regarding health issues, and it was published for the first time in 1999 [21]. DISCERN scores between 4 and 5 points are sorted as "Very High", between 3 and 4 as "High", between 2 and 3 as "Average", between 1 and 2 as "Low", and "Very Low" between 0 and 1. Higher scores in the scale indicate higher levels of quality of the information [21].

GQS evaluates the overall quality of resources that can be found online. Each of the five criteria that can be identified in a video can receive one point, 5 points being the highest educational quality [19]. GQS incorporates the quality of the information and its accessibility, the general information flow, and how helpful it would be for any user [17,22]. The classification scale used was the same as for the DISCERN scale (from "Very Low" to "Very High").

All the videos were also evaluated with the HONCode tool, which was developed by the Health on the Net Foundation [23,24], a nonprofit organization accredited by the United Nations, who elaborated the code of conduct in order to standardize the reliability of online medical/health information [24,25]. HONCode is the most frequently used assessment tool [26] for reliability and credibility of the information that can be found on health-related websites and its certification is submitted to an annual review process made by the HON-Foundation, who also responds to any violation reported by internet users.

The HONCode is not aimed to grade the information quality that is contained in a website, but rather establishes a series of rules in order for website publishers to comply with basic ethical standards of information delivery and to help ensure that visitors are always aware of the purpose and the source of the data they are viewing. It is constituted by a group of parameters about the reliability and credibility of the information that can be found in health-related websites. It is based on 8 principles determining the reliability of web pages and it can score from 0% to 100% [25].

The mentioned principles are as follows: principle 1 is about "*Authority*" and it checks the authors' qualifications; principle 2 checks the "*Complementarity*" regarding the information to support and not to replace; principle 3 is about "*Confidentiality*", regarding the respect the site users' privacy; principle 4 checks the "*Attribution*", i.e., the citation of the dates and sources of the medical information; principle 5 is about "*Justifiability*", that is, the justification of claims and if they are balanced and objective; principle 6 checks the "*Transparency*" regarding accessibility and the delivery of valid contact details; principle 7 is about "*Financial disclosure*", and it checks if funding details are provided. Finally, principle

8 checks the "*Advertising*", i.e., if it distinguishes advertisements in a clear way from the editorial matter [25].

2.3. Statistical Analysis

On the dataset, two different statistical analyses were carried out: The first of them, which was modeled on techniques of machine automated learning, considers a problem of binary classification. It divides, using principal component analysis (PCA) [27], the dataset into two groups: "C1" or "*Relevant Videos*" and "C2" or "*Not Relevant Videos*". With PCA, the main information in a set of data can be visualized and described by multiple and interrelated variables. The information of any dataset is matched with the total variation. So, PCA finds some directions (known as principal components) where the variance is maximized in the data. It is very useful for extracting the most important information from a set of data and it expresses the information referred to the new group of variables (or principal components). With this, PCA shrinks the number of dimensions of a set of data to a smaller number of dimensions or principal components. For this study, two dimensions were enough to bundle the data to be represented graphically, with a minimal forfeiture of the information of the original dataset.

The aim of the first statistical test is to establish which variables distinguish in the best way the groups into which videos are divided. The matter of this classification was checked in three situations according to these variables:

I. DISCERN: Videos classified in the "*C1/Relevant Videos*" group are the samples whose scores were bigger than the mean; the opposite was defining the second group, "C2".
II. Exercises number: Comparable with the preceding variable.
III. PCA: The binary variable defines both groups after a principal component analysis with a k-means algorithm. With this, C1 and C2 groups are defined by information variability with no loss of previous information. Accordingly, no information is lost when the dimensions are reduced to only two PCA, in this case.

For these analyses, a grouping algorithm was used to obtain the least number of dimensions that best divides the videos into two different clusters. Precision was established by *Leave-One-Out Cross Validation* (LOOCV) with the use of the "Nearest Neighbor Classifier". Fisher's ratio (FR) was used to determine the variables' power of discrimination. The higher the Fisher's ratio is, the more discriminatory they are, which is a consequence of their low intraclass dispersion and their high interclass distance. In the case of a binary classification, the FR of a variable "j" is determined by:

$$\frac{(\mu_{j1} - \mu_{j2})^2}{{\sigma_{j1}}^2 + {\sigma_{j2}}^2}$$

where μ_{ji} is the measurement of the mass center of the distribution of probability of the variable "j" in group "i" (i = 1, 2), while σ_{ji} is the measurement of the dispersion inside that group.

Apart from FR, the Spearman, Kendal, and Pearson correlation factors were also calculated for them according to the determined classes. These additional factors disclose the importance, or discrimination power, of these variables according to the classification criteria.

The second statistical analysis used the Wilcoxon test and the t-test, showing the significance of the differences among both groups according to each variable's point of view. This second test determines how relevant the variables are according to the v classification of the videos within group "C1" and group "C2", for all three situations. A t-test establishes whether the mean of a variable in a group has a significant difference with respect to the other. The H0, or null hypothesis, considers no such difference among these groups. If the result of the test is 1, it shows a big enough evidence that it rejects the H0 hypothesis, while if the result is 0, such H0 will be accepted. The accepted level of error probability or statistical significance is $\alpha < 0.05$. Regarding the Wilcoxon test, it determines the difference among every set of couples and it tests this difference. According to this, its H0 or null hypothesis considers the equivalence between population medians for both videos' groups "C1" and "C2".

According to principal components analysis, Henri Kaiser (1970) presented a measure of sampling adequacy (MSA) for factor analytic data matrices that was subsequently adapted by Kaiser and Rice (1974). It is the function of the square of the elements of the matrix when they are compared to the original correlations' squares. This factor, renowned as the Kaiser–Meyer–Olkin (KMO) index, is considered as "Unacceptable" if it is under 0.50, "Miserable" among 0.50 and 0.60, "Mediocre" if it is more than 0.60 and less than 0.70, "Middling" among 0.70 and 0.80, "Meritorious" for 0.80 to 0.90, and if it is more than 0.90 (and less than 1.00) it is classified as "Marvelous" [28]. The KMO test was used to determine the multivariate normality and the sample adequacy. With the objective of evaluating the validity of the construct, the sample suitability for factor analysis was made using Bartlett's test of sphericity [29]. Additionally, the statistical power of the statistical analysis carried out was checked through the combination of the PCA analysis and the KMO index.

DISCERN, HONCode, GQS, and number of exercises, regarding PCA1, show a Pearson correlation factor of 0.867, 0.791, 0.964, and 0.504, respectively, with a p-value < 0.01, which denotes a high statistically significant relation.

For the assessment of the concordance between examiners, the intraclass correlation coefficient (ICC) analysis was performed with confidence intervals (CI) of 95% considering a two-way random model, mean rating (k = 2) consistency, and the method of Pearson's correlation. The significance level was fixed at $p < 0.05$. With a confident interval of 95%, values resulting from the ICC calculation under 0.5 mean a "Poor" reliability, "Moderate" for values between 0.5 and 0.75, "Good" between 0.75 and 0.90, and "Excellent" reliability for values higher than 0.90 [30].

3. Results

Characteristics of Videos

The majority of the analyzed videos were created by non-healthcare workers (47%) and sports institutions (25%). Those produced by healthcare workers and media share a percentage of 9%, respectively, academic institutions reached 6% and, finally, videos produced by health institutions reached 4% of the total. As for the origin, 63% of the videos were made in America, followed by Europe (19%), Asia (15%) and, finally, Africa and Australia with 1% each.

The target audiences for which the videos are intended are mostly adults (80.9%). We found that videos intended for children reached only 1.5% of the total, 5.9% for the elderly, and 11.8% for all audiences. Of the total, only 1.5% of the videos are intended for users with some type of pathology. To carry out the exercises, the authors do not use any type of material in 39.7% of the videos, in 29.4% of them some kind of domestic material is used, and in 30.9% of the cases the author uses professional material. The data collected regarding the exercises' time show that in 75.0% of the videos, they did not adjust to the time recommended by WHO for this type of activity, compared to 25.0% of them that did adjust to it. The inter-reviewer agreement for this study scored 0.9299, which points to the concordance between both examiners for a categorization as "Excellent". The Kaiser–Meyer–Olkin measure for the adequacy of the sample (KMO index) for the PCA was performed to check the sample size compliance before applying the analysis of factors. The KMO measure is a test that is used to decide whether a group of samples are suitable for conducting factor analysis, and it is calculated in terms of the correlation and partial correlation between the variables. The result of this test showed that the KMO value reached 0.845. According to this, it can be determined that the size of the sample was, according to Kaiser, "Meritorious" for this dataset. This KMO measure also indicates the power of statistical analysis that was carried out.

The sphericity test of Bartlett (BST) [29] was performed to determine if there is an intra-correlation between the dataset variables considering partial correlations. It unveiled that there was a statistically significant relation ($\chi^2 = 142.577$; $p < 0.001$), so the sample had a good adequacy for factorized analysis, the data shows a good appropriateness for statistical-related assumptions according to multivariated normality and the data were not comprising an identity matrix. So, it can be concluded that, due to the fact that the calculated X^2-test resulted as significant, the data matrix could be determined as appropriate. The applicability of the construct was confirmed due to the fat that the varimax rotation method and the extraction method were combined in the PCA analysis. The total variance explained with this PCA analysis was 85.507%.

The data used for each video were the mean of the scores assigned by each of the independent researchers, including the HONCode, GQS, and DISCERN scores. According to these scores, the DISCERN mean was 2.29 (SD 0.71), HONCode scored a mean of 58.95 (SD 12.89), and GQS had a mean of 2.32 (SD 0.86). Descriptive statistic basic data of the mean results for HONCode, GQS, and DISCERN are shown in Table 1.

Table 1. Descriptive statistics of the videos.

Descriptive Statistics	
(Mean ± Standard Deviation)	
Video length (minutes)	10:37 ± 10:55
View count (*n*)	59,565 ± 229,286
Days online (*n*)	42.3 ± 13.6
View ratio (Views/day)	1326.62 ± 4860.72
Likes (*n*)	2275 ± 9575
Dislikes (*n*)	29 ± 105
Likes ratio (likes/day)	50.75 ± 205.78
Dislikes ratio (dislikes/day)	0.67 ± 2.29
VPI (%)	97.33 ± 3.66
Subscribers (*n*)	238,278 ± 680,369
DISCERN score	2.29 ± 0.71
HON score	58.95 ± 12.89
Global quality score	2.32 ± 0.86

Regarding the scores of the videos in the different questions that DISCERN applies (considering mean ± SD), it is remarkable that question 1 scored highest (0.93 ± 0.17); followed by question 3 (0.77 ± 0.36) and, then, question 5 (0.25 ± 0.32). The lowest scoring questions were question 4 (0.24 ± 0.33) and question 2 (0.10 ± 0.28). Overall results were 2.29 ± 0.91 with a median of 2.25 (min 0.50; max 4.50).

Regarding the mean scores for DISCERN, videos' reliability performed in 27.9% of the cases as "Very Poor", in 45.6% of them as "Poor", 17.6% as "Average", 8.8% as "High", and there was no video with a "Very High" score. However, referring to the mean GQS, the videos' quality performed in 22.1% of the cases as "Very Poor", in 54.4% as "Poor", in 14.7% as "Average", in 8.8% as "High", and no "Very High" videos were found.

The coefficients of Pearson's correlation for all variable's referred to DISCERN score were also calculated. These calculations show the discrimination power (relevance) of the rest of the indicators considering a level of significance of <0.05 (Table 2).

Table 2. Coefficients of Pearson's correlations of each variable's referred to DISCERN score.

DISCERN *	Pearson Coefficient	Std Error	t	$p-$Value **	Lower	Upper
VPI	−0.1593	0.1215	−13.108	0.1945	−0.4019	0.0833
Views per day	0.2296	0.1198	19.168	0.0596	−0.0096	0.4688
Likes per day	0.2171	0.1202	18.069	0.0753	−0.0228	0.4570
Dislikes per day	0.2234	0.1200	18.615	0.0671	−0.0162	0.4629
GQS	0.9158	0.0494	185.266	<0.001	0.8171	10.145
HONCode	0.6055	0.0980	61.813	<0.001	0.4099	0.8011
Exercises	0.5397	0.1036	52.077	<0.001	0.3328	0.7466

* Pearson's correlation coefficients (t-test—2 tailed). ** $\alpha = 0.05$.

Regarding the relation between the author and the videos' reliability, considering DISCERN scores, videos produced by healthcare workers show the highest scores (3.58), followed by academic institutions (3.25), health institutions (2.83), media (2.50), sports institutions (2.41) and, with the lowest score, non-healthcare workers (1.78). Similarly, according to the other variables referring to educational quality, GQS, and videos whose authors are healthcare workers also showed higher scores (3.50), followed by health institutions (3.33), academic institutions (3.63), sport institutions (2.44), media (2.42) and, finally, non-healthcare workers (1.84). This data (with the corresponding SD of each of them) is showed in Table 3 with the views ratio scores too. It also can be seen in Table 3 that the top scores in GQS and DISCERN scales were found in videos whose origin was Australia (3.50 in both cases), followed by those produced in Europe (2.58 and 2.54, respectively), Asia (2.25 and 2.30), America (2.20 and 2.24), and Africa (2.00 for both scales).

Table 3. Author and origin variables related with DISCERN, GQS, HONCode, and views ratio.

		DISCERN		GQS		HONCode		Views Ratio	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Source	Academic institution	3.25	0.82	3.00	1.00	66.63	19.70	280.39	156.30
	Sports institution	2.41	0.69	2.44	0.56	60.62	15.49	328.83	834.05
	Media (newspaper, TV)	2.50	0.52	2.42	0.00	58.50	11.90	2164.23	3171.15
	Health institution	2.83	0.58	3.33	1.00	72.33	12.17	272.30	343.12
	Non-healthcare workers	1.78	0.53	1.84	0.69	53.06	15.23	2172.89	6861.22
	Healthcare workers	3.58	0.82	3.50	0.82	74.25	14.51	27.31	21.07
	TOTAL	2.29	0.71	2.32	0.86	58.95	12.89	1326.62	4860.72
Origin	Africa	2.00	0.00	2.00	0.00	63.00	0.00	1907.49	0.00
	America	2.20	0.73	2.24	0.80	56.80	18.00	1532.59	5708.68
	Asia	2.25	0.48	2.30	0.74	61.85	10.15	825.62	2414.64
	Australia	3.50	0.00	3.50	0.00	69.00	0.00	60.24	0.00
	Europe	2.58	0.97	2.54	1.03	62.73	20.91	1083.46	3590.75
	TOTAL	2.29	0.71	2.32	0.86	58.95	12.89	1326.62	4860.72

Regarding the five classes in which DISCERN distinguishes the videos, their VPI scores were as follows: videos that scored "Very Poor" in DISCERN had a VPI score of 97.37, videos that scored "Poor" in DISCERN had a VPI of 97.96, videos that scored "Average" in DISCERN had a VPI score of 96.53, videos that were "High" in the DISCERN score had a VPI score of 95.61, and there were no videos classified as "Very High" for DISCERN. No correlation with statistical significance was found between GQS-VPI and between DISCERN-VPI scores ($p = 0.194 > 0.05$ for DISCERN and VPI; $p = 0.270 > 0.05$ for GQS and VPI).

Referring to the number of exercises coinciding with those recommended by WHO (12 exercises), no video contains more than 5 exercises; out of those 12, only 1% of them contain 4 and 5 exercises, respectively, 25% contain 3 exercises, 26% contain 2, 31% contain 1 exercise, and 15% of the videos do not contain any of the recommended exercises. That is, the videos that contain two or less exercises represent 72% of the total and the videos that contain more than two exercises represent 28% of the total. Videos with two or less exercises obtain a DISCERN mean of 2.02, a HONCode mean of 55.52, GQS mean 2.08, views ratio mean of 685.43, likes ratio mean of 22.23, dislike ratio mean of 0.30, and a VPI mean of 98.06. Videos that include more than two exercises receive a DISCERN mean of 3.00, HONCode mean of 67.79, GQS mean of 2.95, views ratio mean of 2980.21, likes ratio mean of 124.32, dislike ratio mean of 1.62, and VPI mean of 95.46 (Table 4).

Table 4. Descriptive statistics referred to the number of exercises shown in the videos.

		Number of Exercises		
		≤2 (0-1-2)	**>2 (3-4-5)**	**Total**
n (%)		72.06	27.94	100%
DISCERN	Mean	2.02	3.00	2.29
	SD	0.65	0.83	0.71
HONCode	Mean	55.52	67.79	58.95
	SD	16.68	10.33	12.89
GQS	Mean	2.08	2.95	2.32
	SD	0.72	0.92	0.86
Views ratio	Mean	685.43	2980.21	1326.62
	SD	2751.60	8156.53	4860.72
Likes ratio	Mean	22.23	124.32	50.75
	SD	100.99	358.01	205.78
Dislikes ratio	Mean	0.30	1.62	0.67
	SD	1.23	3.85	2.29
VPI	Mean	98.06	95.46	97.33
	SD	2.78	4.90	3.66

Regarding the PCA outlook of the statistical assessment, if it is considered a graphical representation of the coordinates of each video in a coordinate system where the axes are the principal components, it can be seen that videos are grouped naturally and clearly (Figure 2), with no bias or manual intervention, in two clusters of points. These two clusters include those videos with higher scores in GQS, number of exercises, DISCERN and HONCode and in one cluster (C1), and videos with lower scores in the other cluster (C2) (Figure 2).

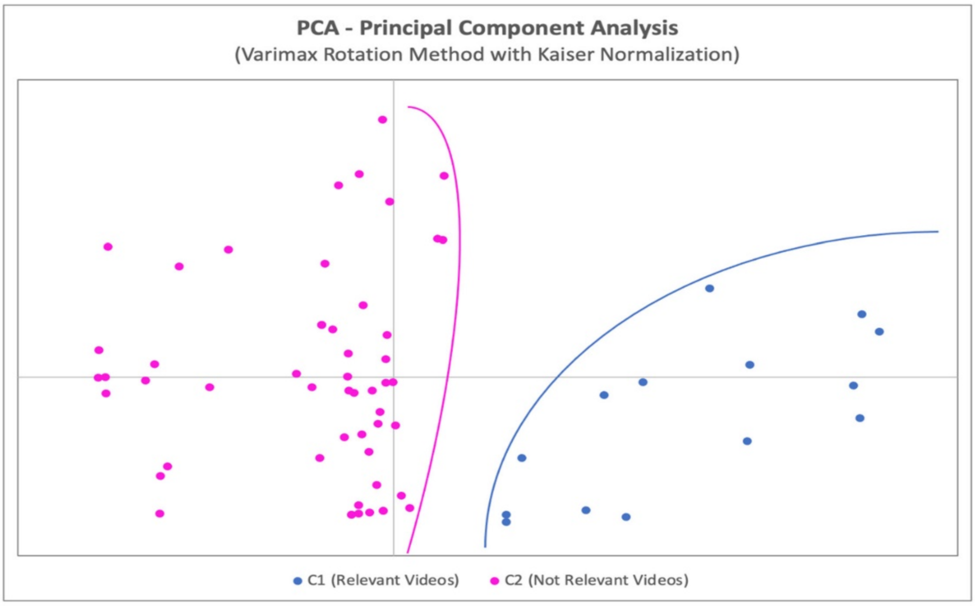


Figure 2. Principal component analysis graphical auto-classification of the videos.

4. Discussion

The objective of this study was to evaluate the quality of YouTube videos related to home exercises during lockdown and their adherence to WHO recommendations. According to this, the main findings of the present research is that the existing videos on the YouTube platform related to PA during lockdown are of low quality and lack concordance with WHO recommendations.

Importantly, based on the data obtained, it seems that the videos related to PA during lockdown available on this platform are not produced by professionals, nor do they present reliable information since the majority of them were developed by non-healthcare workers (47%), while the academic institutions and healthcare workers only constituted 6% in both cases. This information is consistent with other research indicating that these health institutions are underrepresented in publishing videos related to health information [16] and reporting that the majority of the examined YouTube videos about breast self-examination are uploaded by individual users [22]. Regarding to DISCERN ratings, reliability and quality of most videos (73.5%) was found to be "Very Poor" (27.9%) and "Poor" (45.6%). Similar scores were obtained in the majority of videos (76.5%) for GQS (22.1% and 54.4%, respectively). There is a relation with statistical significance between GQS and DISCERN measures (Table 2). The present results highlight that caution should be taken when consuming online exercise classes in spite of the encouragement provided by the WHO in this regard.

HONCode is related to the ethics of the information presented and also with the information quality and reliability [26], which agrees with our study, that shows a relation with statistical significance between HONCode and DISCERN measures (Table 2). Additionally, videos uploaded by healthcare workers presented higher scores in GQS, HONCode, and DISCERN scales, followed by videos made by academic institutions and health institutions. These data are consistent with other studies indicating that the YouTube channels of universities produced those videos with the most precise information [31]. The lowest figures in terms of educational quality and reliability are those videos made by non-healthcare workers, which are, on the other hand, the most numerous (Table 3).

The views ratio shows the highest figures in videos that scored "High" in the DISCERN scale, which indicates that users mostly view high-quality videos (despite the fact that the quantity of these videos is low). Similar information was reported by the study of Sood et al. where helpful videos have many more visitors than deceptive videos [31]. Those videos with "Poor" quality score in the DISCERN scale achieve the highest scores in video popularity—VPI (97.96%)—while videos with "High" DISCERN score receive the lowest VPI (95.61%). These data agree with another study where videos that obtained the highest VPI scores had the lowest reliability values [32]. Moreover, according to another study, no correlation with statistical significance was identified between DISCERN, GQS values, and VPI scores [12].

Regarding the length of the videos, the results of this study agree with the literature; research videos averaged 10:37 min and preceding research found length means from 6:17 to 10:35 min [33]. Interestingly, only 25% of the videos follow the WHO recommendations regarding weekly exercise time, while 75% of the videos do not. In addition, in 39.7% of the videos, the exercises are performed without material, and in 29.4% domestic material is used. In the remaining 30.9% of the videos, the user requires some kind of technical material to carry out the exercises that does not fit with the WHO recommendations.

Analyzing educational quality and reliability, videos containing two or more recommended exercises show higher scores in DISCERN and GQS scales than videos containing two or fewer exercises, despite these being the most numerous (72% of the total). The same occurs in terms of the activity duration, since the videos that coincide in with WHO recommendations for exercise length are the best valued in these scales. That is, the videos that best fit the recommended characteristics in exercise type and completion time are the ones that obtain the best quality and reliability scores.

WHO recommends a series of 12 exercises, but none of the analyzed videos contain more than 5 of those. Of the total, 31% of the videos contain one exercise, 26% contain two exercises, 25% contain three, 15% of the videos do not indicate any of the recommended exercises, 1% contain four exercises, and another 1% contain five exercises. These figures show little coincidence between the recommendations of an organization such as WHO and the information that is available at user level in the platform that this organization specifically recommends on its website. According to Jamal et al., who systematically reviewed the effect of health information systems (HIS) on health care quality and did not find enough evidence of clinical or statistical relevant improvements in patient outcomes [34], recommendations such as those made by WHO don't achieve the expected improvements in patients because, in this case, these recommendations are not delivered to users. Since international health organizations such as WHO recommend a series of home exercises for quarantine periods to keep people's health in a good situation and they suggest that people can find videos on YouTube where these exercises are explained, these organizations, such as WHO in this case, should verify whether their recommendations are met on the video platforms on internet where they suggest that their recommended exercises are better explained. These recommendations are focused on being physically proactive, avoiding inactive conduct and levels of PA similar to a sedentary lifestyle. This kind of behavior could lead to adverse impacts on the health, quality of life, and well-being of people. Since WHO prepared a list of home exercises as recommendations for people in lockdown, and in these recommendations it is mentioned that they can be found on online exercises classes where many of them are free, such as YouTube, WHO should stablish some kind of control to check if their own advice is met on the platforms that they specifically mention.

Biomedical institutions and public organizations should consider if YouTube supplies viewer and (possible) patients with precise and useful data, or if its videos' contents are possibly damaging and deceptive. These data agree with other authors stating that the data found in YouTube videos often contradict the medical recommendations and standards [35]. According to this, medical organizations and authorities should review the relevance and precision of the information that can be found on internet and should offer their help to society in accessing the most precise information [12]. All this data should be institutionally filtered. In fact, in 2001, Loretti et al. published that the WHO must improve its own performance [36], and 19 years later, with the development of eHealth and the use of the internet as one of the most broad information suppliers used, this need is even more important. The strategies necessary to develop reliable online information bring about the need to create tools that assess its quality and, although some were already developed by different scientific associations and organizations such as the American Medical Association (AMA), the Internet HealthCare Coalition, or the Health-On-the-Net Foundation (HON) [32], YouTube does not apply them despite the fact that many researchers suggest that confidence-worthy medical institutions and organizations should have more presence on YouTube and supply useful data that patients can trust [35].

After performing these two statistical analyses, the PCA calculation with the KMO index allows the videos to auto-classify into two clusters (C1 and C2, high and low quality videos, respectively) without losing any information of any variable of the study. All the variables are combined, through the PCA methodology, into two new variables that contain all the information of the initial variables. According to this, the quality of the videos can be assessed attending only to these new variables, which makes the assessment much quicker, easier, and with no lost of information, which represents an enormous advantage when performing this kind of analyses, also because the *p*-values show a relationship with high statistical significance.

The statistical strength of these two analyses is so great that it permits the self-classification of the videos, considering all their characteristics (quality among them), using only two variables (the principal components) which incorporate the data of the totality of the variables considered in this research.

Limitations

In order to simulate standard user behavior, searches were not conducted as incognito in order to prevent the effects of geographic location and browsing history that could limit the results because some users/viewers can make this kind of modification.

This research shows the status of YouTube content at a given moment [32], due to the fact that it is a website that evolves constantly. Furthermore, the results of this study can only be considered for this website since it only takes this one into account, as mentioned in the selection criteria.

Additionally, because the study tried to reproduce the typical user's search behavior [32], only the first 150 results were considered due to the fact that most internet visitors do not search beyond the first 50 results.

In this research, no viewers/consumer characteristics or intentions were evaluated when watching the videos. Furthermore, no video comments were considered from the comments section that YouTube adheres to each video.

5. Conclusions

The quality of available videos in YouTube concerning PA during lockdown is low and does not reflect WHO's recommendations. Organizations should consider if YouTube supplies viewers and (possible) patients with precise and useful data or if its videos' contents are possibly damaging and deceptive. Effective strategies and policies capable of indicating the quality of this information are needed to filter out erroneous or non-rigorous information that may affect people's health. These tools should help any user/viewer to distinguish videos of high and low quality.

Author Contributions: Conceptualization, A.M.R.-R., M.B.-D. and S.H.-S.; data curation, M.d.l.F.-C. and J.C.; formal analysis, A.M.R.-R.; investigation, M.B.-D., J.C. and I.E.-P.; resources, M.d.l.F.-C.; software, A.M.R.-R.; supervision, A.M.R.-R., M.B.-D. and S.H.-S.; visualization, I.E.-P.; writing—original draft, M.B.-D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

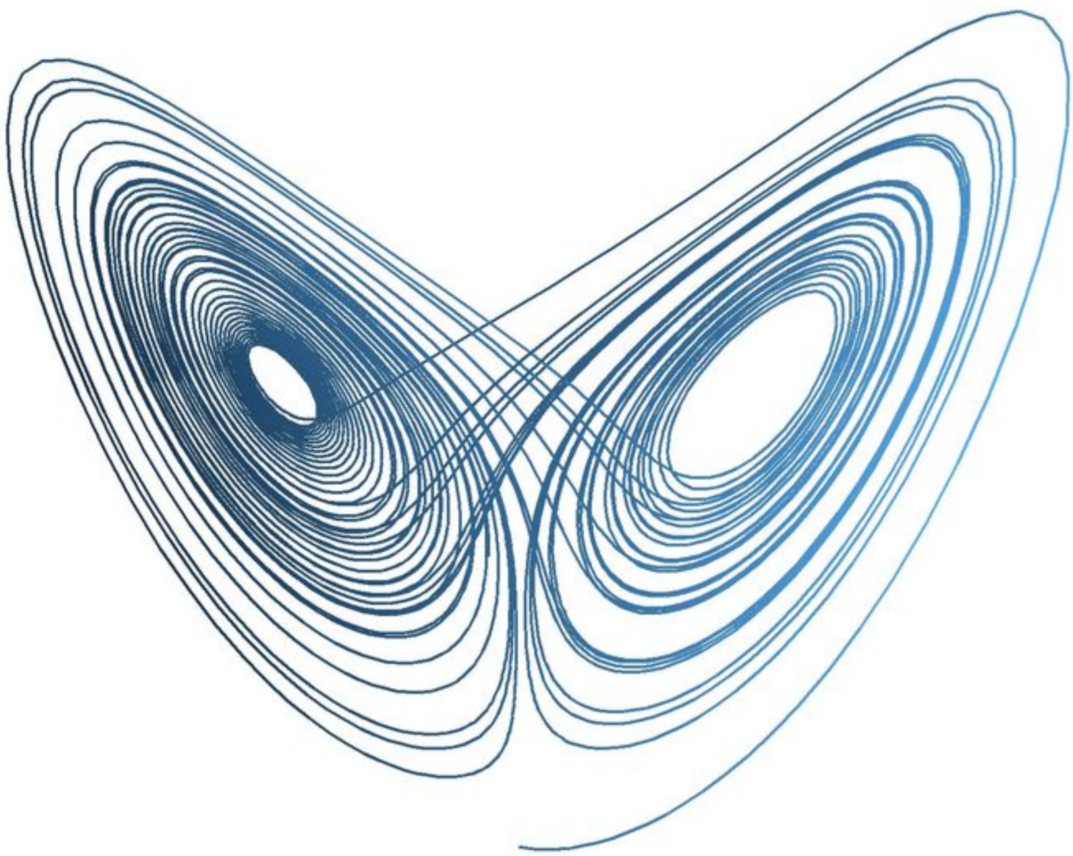
Data Availability Statement: All data generated or analyzed during this study are included in this article. Further enquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bulut, C.; Kato, Y. Epidemiology of COVID-19. *Turk. J. Med. Sci.* **2020**, *50*, 563–570. [CrossRef]
2. Shah, S.G.S.; Farrow, A. World Health Organization declares global emergency: A review of the 2019 novel Coronavirus (COVID-19). *Int. J. Surg.* **2020**, *76*, 128–129. [CrossRef] [PubMed]
3. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard with Vaccination Data. Available online: https://covid19.who.int/ (accessed on 17 April 2022).
4. World Health Organization. *Stay Physically Active during Self-Quarantine*; World Health Organization: Geneva, Switzerland, 2020; pp. 1–7.
5. World Health Organization. Physical activity. *Prim. Health. Care* **2003**, *13*, 8. [CrossRef]
6. Burtscher, J.; Burtscher, M.; Millet, G.P. (Indoor) Isolation, Stress and Physical Inactivity: Vicious Circles Accelerated by COVID-19? *Scand. J. Med. Sci. Sports* **2020**, 13706. [CrossRef] [PubMed]
7. Zieff, G.; Bates, L.C.; Kerr, Z.Y.; Moore, J.B.; Hanson, E.D.; Battaglini, C.; Stoner, L. Targeting sedentary behavior as a feasible health strategy during COVID-19. *Transl. Behav. Med.* **2020**, *11*, 826–831. [CrossRef]
8. World Health Organization. Recomendaciones Mundiales Sobre Actividad Física Para La Salud; 2010. Available online: https://apps.who.int/iris/bitstream/handle/10665/44441/9789243599977_spa.pdf;jsessionid=F0D75147BCDB443B786DAADB4195C6D1?sequence=1 (accessed on 1 November 2020).
9. Matias, T.; Dominski, F.H.; Marks, D.F. Human needs in COVID-19 isolation. *J. Health Psychol.* **2020**, *25*, 7. [CrossRef]
10. Chen, P.; Mao, L.; Nassis, G.P.; Harmer, P.; Ainsworth, B.E.; Li, F. Wuhan coronavirus (2019-nCoV): The need to maintain regular physical activity while taking precautions. *J. Sport Health Sci.* **2020**, *9*, 103–104. [CrossRef]
11. World Internet Users Statistics and 2020 World Population Stats. Copenhagen Business School ISIS MISC. Publ. Available online: https://www.coursehero.com/file/58641086/World-Internet-Users-Statistics-and-2020-World-Population-Statspdf/ (accessed on 7 November 2020).
12. Atci, A.G. Quality and reliability of the information on YouTube Videos about Botox injection on spasticity. *Rom. Neurosurg.* **2019**, *33*, 473–477. [CrossRef]
13. Omnicore Agency. YouTube by the Numbers (2022): Stats, Demographics & Fun Facts. Available online: https://www.omnicoreagency.com/youtube-statistics/ (accessed on 17 April 2022).
14. Aydin, M.A.; Akyol, H. Quality of Information Available on YouTube Videos Pertaining to Thyroid Cancer. *J. Cancer Educ.* **2019**, *18*, 1–14. [CrossRef]
15. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med.* **2009**, *6*, e1000097. [CrossRef]
16. Kuru, T.; Erken, H.Y. Evaluation of the Quality and Reliability of YouTube Videos on Rotator Cuff Tears. *Cureus* **2020**, *12*, e6852. [CrossRef]
17. Erdem, M.N.; Karaca, S. Evaluating the accuracy and quality of the information in kyphosis videos shared on youtube. *Spine (Phila Pa 1976)* **2018**, *43*, E1334–E1339. [CrossRef]
18. Charnock, D. *The Discern Handbook*; Radcliffe Medical Press: Abingdon, UK; Oxford, UK, 1998.
19. Kunze, K.N.; Cohn, M.R.; Wakefield, C.; Hamati, F.; LaPrade, R.F.; Forsythe, B.; Yanke, A.B.; Chahla, J. YouTube as a Source of Information About the Posterior Cruciate Ligament: A Content-Quality and Reliability Analysis. *Arthrosc. Sports Med. Rehabil.* **2019**, *1*, e109–e114. [CrossRef]
20. Singh, A.G.; Singh, S.; Singh, P.P. YouTube for information on rheumatoid arthritis—A wakeup call? *J. Rheumatol.* **2012**, *39*, 899–903. [CrossRef]
21. Charnock, D.; Shepperd, S.; Needham, G.; Gann, R. DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *J. Epidemiol. Community Health* **1999**, *53*, 105–111. [CrossRef]
22. Esen, E.; Aslan, M.; Sonbahar, B.Ç.; Kerimoğlu, R.S. YouTube English videos as a source of information on breast self-examination. *Breast Cancer Res. Treat.* **2019**, *173*, 629–635. [CrossRef]
23. HONCode: E-Guide for Health Consumers. Available online: https://www.hon.ch/HONCode/Pro/HealthEvaluationTool.html (accessed on 26 April 2020).
24. Boyer, C.; Selby, M.; Scherrer, J.R.; Appel, R.D. The health on the Net Code of Conduct for medical and health Websites. *Comput. Biol. Med.* **1998**, *28*, 603–610. [CrossRef]
25. Boyer, C. When the Quality of Health Information Matters: Health on the Net in the Quality Standard for Information You Can Trust. 2013. Available online: https://www.hon.ch/Global/pdf/TrustworthyOct2006.pdf (accessed on 26 April 2020).
26. Fahy, E.; Hardikar, R.; Fox, A.; Mackay, S. Quality of patient health information on the internet: Reviewing a complex and evolving landscape. *Australas. Med. J.* **2014**, *7*, 24–28. [CrossRef]
27. Zhang, Z.; Castelló, A. Principal components analysis in clinical studies. *Ann. Transl. Med.* **2017**, *5*, 351. [CrossRef]
28. Kaiser, H.F. Educational and Psychological Measurement. *Ser. Rev.* **1982**, *8*, 83–85.

29. Bartlett, M.S. The Effect of Standardization on a X 2 Approximation in Factor Analysis. *Biometrika* **1951**, *38*, 337–344. [CrossRef]
30. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef] [PubMed]
31. Sood, A.; Sarangi, S.; Pandey, A.; Murugiah, K. YouTube as a source of information on kidney stone disease. *Urology* **2011**, *77*, 558–562. [CrossRef] [PubMed]
32. Radonjic, A.; Fat Hing, N.N.; Harlock, J.; Naji, F. YouTube as a source of patient information for abdominal aortic aneurysms. *J. Vasc. Surg.* **2020**, *71*, 637–644. [CrossRef]
33. Gokcen, H.B.; Gumussuyu, G. A Quality Analysis of Disc Herniation Videos on YouTube. *World Neurosurg.* **2019**, *124*, 799–804. [CrossRef]
34. Jamal, A.; McKenzie, K.; Clark, M. The impact of health information technology on the quality of medical and health care: A systematic review. *Health Inf. Manag. J.* **2009**, *38*, 26–37. [CrossRef]
35. Allgaier, J. Science and Medicine on YouTube. In *Second International Handbook of Internet Research*; Hunsinger, J., Allen, M.M., Klastrup, L., Eds.; Springer: Dordrecht, The Netherlands, 2020; pp. 7–27. [CrossRef]
36. Loretti, A.; Leus, X.; Van Holsteijn, B. Relevant in Times of Turmoil: WHO and Public Health in Unstable Situations. *Prehosp. Disaster Med.* **2001**, *16*, 184–191. [CrossRef]



Lorenz´s attractor for deterministic nonlinear dynamic systems figure
(Butterfly effect in the Chaos Theory)

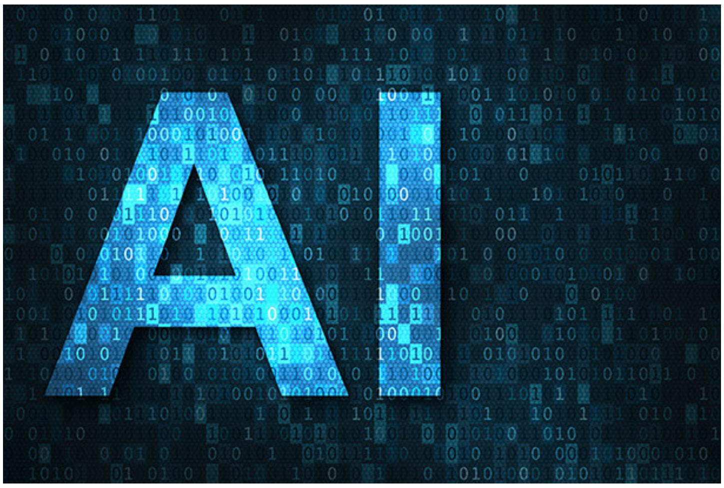
"In the midst of chaos, there is also opportunity"
Sun-Tzu

9

9. Tables and Figures

9.1. Tables

9.2. Figures



Logo used by UNESCO on November 24th 2021, when the *Recommendation on the Ethics of Artificial Intelligence* was adopted by UNESCO's General Conference at its 41st session.

"Ethics and Science need to shake hands"

Richard Clarke Cabot

10

10. Ethics Commission approval

The Ethics Committee of Research in Humans of the Ethics Commission in Experimental Research of University of Valencia,

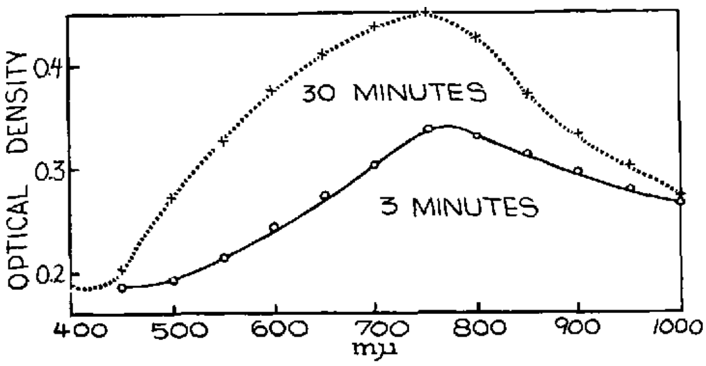
CERTIFY:

Hereby certify that the Ethics Committee of Research in Humans, in the session which took place on , analysed the project of doctoral thesis entitled *"Quality analysis of available information for patients with different pathologies on social media video platforms."*, whose researcher in charge is JOSÉ CASAÑA GRANELL, and agreed with this project.

And in witness whereof, I hereby sign this certificate

Av. Blasco Ibáñez, 13 tel: 963864109 vicerec.investigacio@uv.es
València 46010 fax: 963983221 www.uv.es/serinves

Firmado digitalmente por
PEDRO JESUS PEREZ ZAFRILLA
Cargo: Presidente del Comité de Ética de la Investigación en Humanos
Fecha: 12/09/2021 23:17:33 CEST



Absorption spectra graph presented in the most cited article of all time

(*Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. J Biol Chem. 1951 Nov;193(1):265-75. PMID: 14907713.*)

"It is not necessary to be a genius to contribute to science"

Oliver H. Lowry

11

11. References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459. https://doi.org/10.1002/WICS.101

Afful-Dadzie, E., Afful-Dadzie, A., & Egala, S. B. (2021). Social media in health communication: A literature review of information quality. *Health Information Management : Journal of the Health Information Management Association of Australia*. https://doi.org/10.1177/1833358321992683

Ahmad, T., Sattar, K., & Akram, A. (2020). Medical professionalism videos on YouTube: Content exploration and appraisal of user engagement. *Saudi Journal of Biological Sciences*, *27*(9), 2287–2292. https://doi.org/10.1016/J.SJBS.2020.06.007

Akyol, A., & Karahan, İ. (2020). Is YouTube a quality source of information on sarcopenia? *European Geriatric Medicine*, *11*(4), 693–697. https://doi.org/10.1007/S41999-020-00327-W

Alefishat, E., Farha, R. A., & Zawiah, M. (2021). Pharmacists' perceptions of the use of internet-based medication information by patients: A cross-sectional survey. *PLoS ONE*, *16*(8). https://doi.org/10.1371/JOURNAL.PONE.0256031

Altan Şallı, G., Egil, E., Şallı, G. A., & Egil, E. (2020). Are YouTube videos useful as a source of information for oral care of leukemia patients? *Quintessence International*, *51*(1), 78–85. https://doi.org/10.3290/j.qi.a43665

Amante, D. J., Hogan, T. P., Pagoto, S. L., English, T. M., & Lapane, K. L. (2015). Access to Care and Use of the Internet to Search for Health Information: Results From the US National Health Interview Survey. *J Med Internet Res 2015;17(4):E106 Https://Www.Jmir.Org/2015/4/E106*, *17*(4), e4126. https://doi.org/10.2196/JMIR.4126

American Cancer Society. (2019). *Breast Cancer: Treating Breast Cancer*. American Cancer Society.

American Cancer Society. (2021). *Exercises After Breast Cancer Surgery*. https://www.cancer.org/cancer/breast-cancer/treatment/surgery-for-breast-cancer/exercises-after-breast-cancer-surgery.html

Amezcua-Prieto, C., Fernández-Luna, J. M., Huete-Guadix, J. F., Bueno-Cavanillas, A., & Khan, K. S. (2020). Artificial intelligence and automation of systematic reviews in women's health. *Current Opinion in Obstetrics & Gynecology*, *32*(5), 335–341.

https://doi.org/10.1097/GCO.0000000000000643

Ashtari, S., & Taylor, A. D. (2022). The Internet Knows More Than My Physician: Qualitative Interview Study of People With Rare Diseases and How They Use Online Support Groups. *Journal of Medical Internet Research*, *24*(8), e39172. https://doi.org/10.2196/39172

Atci, A. G. (2019). Quality and reliability of the information on YouTube Videos about Botox injection on spasticity. *Romanian Neurosurgery*, *December*, 473–477. https://doi.org/10.33962/roneuro-2019-075

Aydin, M. A., & Akyol, H. (2020). Quality of Information Available on YouTube Videos Pertaining to Thyroid Cancer. *Journal of Cancer Education : The Official Journal of the American Association for Cancer Education*, *35*(3), 599–605. https://doi.org/10.1007/s13187-019-01502-9

Aydın Sayılan, A., & Özbaş, A. (2018). The Effect of Pelvic Floor Muscle Training On Incontinence Problems After Radical Prostatectomy. *American Journal of Men's Health*, *12*(4), 1007–1015. https://doi.org/10.1177/1557988318757242

Azak, M., Şahin, K., Korkmaz, N., & Yıldız, S. (2022). YouTube as a source of information about COVID-19 for children: Content quality, reliability, and audience participation analysis. *Journal of Pediatric Nursing*, *62*, e32–e38. https://doi.org/10.1016/J.PEDN.2021.06.024

Bahar-Ozdemir, Y., Ozsoy-Unubol, T., & Akyuz, G. (2021). Is YouTube a high-quality source of information on cancer rehabilitation? *Journal of Cancer Survivorship*, *16*(5). https://doi.org/10.1007/s11764-021-01093-9

Baker, J. D., Baig, Y., Siyaji, Z. K., Hornung, A. L., Zavras, A. G., Mallow, G. M., Zbeidi, S., Shepard, N. A., & Sayari, A. J. (2021). Assessing the Quality and Credibility of Publicly Available Videos on Cervical Fusion: Is YouTube a Reliable Educational Tool? *International Journal of Spine Surgery*, *15*(4), 669–675. https://doi.org/10.14444/8088

Baran, C., & Yilmaz Baran, S. (2021). Youtube videos as an information source about urinary incontinence. *Journal of Gynecology Obstetrics and Human Reproduction*, *50*(10). https://doi.org/10.1016/J.JOGOH.2021.102197

Bartholomew, D. J. (1995). Spearman and the origin and development of factor analysis. *British Journal of Mathematical and Statistical Psychology*, *48*(2), 211–220. https://doi.org/10.1111/j.2044-8317.1995.tb01060.x

Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). The origin of factor scores: Spearman, Thomson and Bartlett. *British Journal of Mathematical and Statistical Psychology*, *62*(3), 569–582. https://doi.org/10.1348/000711008X365676

Bartlett, M. S. (1951). The Effect of Standardization on a Χ 2 Approximation in Factor Analysis. *Biometrika*, *38*(3–4), 337–344. https://doi.org/10.1093/biomet/38.3-4.337

Bauer, G. R., & Lizotte, D. J. (2021). Artificial Intelligence, Intersectionality, and the Future of Public Health. *American Journal of Public Health*, *111*(1), 98–100. https://doi.org/10.2105/AJPH.2020.306006

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. In *JAMA - Journal of the American Medical Association* (Vol. 319, Issue 13, pp. 1317–1318). JAMA. https://doi.org/10.1001/jama.2017.18391

Ben Salem, K., & Ben Abdelaziz, A. (2021). Principal component analysis (PCA). *Tunisie Medicale*, *99*(4), 383–389. https://doi.org/10.1515/9783110629453-078

Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *International Journal of Environmental Research and Public Health*, *15*(12). https://doi.org/10.3390/IJERPH15122796

Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika*, *65*(1), 23–28. https://doi.org/10.1007/BF02294183

Boyer, C. (2013). *When the quality of health information matters: Health on the Net in the quality Standard for Information You can Trust*. https://www.hon.ch/Global/pdf/TrustworthyOct2006.pdf

Boyer, C., Selby, M., Scherrer, J. R., & Appel, R. D. (1998). The health on the Net Code of Conduct for medical and health Websites. *Computers in Biology and Medicine*, *28*(5), 603–610. https://doi.org/10.1016/S0010-4825(98)00037-7

Braczynski, A. K., Ganse, B., Ridwan, S., Schlenstedt, C., Schulz, J. B., & Hoog Antink, C. (2021). YouTube Videos on Parkinson's Disease are a Relevant Source of Patient Information. *Journal of Parkinson's Disease*, *11*(2), 833–842. https://doi.org/10.3233/JPD-202513

Brar, J., Ferdous, M., Abedin, T., & Turin, T. C. (2021). Online Information for Colorectal Cancer Screening: A Content Analysis of YouTube Videos. *Journal of Cancer Education : The Official Journal of the American Association for Cancer Education*, *36*(4), 826–831. https://doi.org/10.1007/s13187-020-01710-8

Brar, J., Khalid, A., Ferdous, M., Abedin, T., & Turin, T. C. (2022). Breast cancer screening literacy information on online platforms: A content analysis of YouTube videos. *Breast Disease*, *41*(1), 81–87. https://doi.org/10.3233/BD-201028

Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods*, *6*(9), 2812–2831. https://doi.org/10.1039/C3AY41907J

Bulut, C., & Kato, Y. (2020). Epidemiology of covid-19. *Turkish Journal of Medical Sciences*, *50*(SI-1), 563–570. https://doi.org/10.3906/sag-2004-172

Burtscher, J., Burtscher, M., & Millet, G. P. (2020). (Indoor) isolation, stress and physical inactivity: vicious circles accelerated by Covid-19? *Scandinavian Journal of Medicine & Science in Sports*, sms.13706. https://doi.org/10.1111/sms.13706

Cakmak, G., & Mantoglu, B. (2021). Reliability and Quality of YouTube Contents Pertaining to Pancreatic Cancer. *Cureus*, *13*(3), e14085. https://doi.org/10.7759/cureus.14085

Cancer Research UK. (2020). Exercises after breast cancer surgery. *American Cancer Society.* https://www.cancer.org/cancer/breast-cancer/treatment/surgery-for-breast-cancer/exercises-after-breast-cancer-surgery.html%0Ahttp://www.cancerresearchuk.org/cancer-help/type/breast-cancer/treatment/surgery/after-your-breast-cancer-operation

Celebi, M. E., Kingravi, H. A., & Vela, P. A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*, *40*(1), 200–210. https://doi.org/10.1016/J.ESWA.2012.07.021

Chang, M. C., & Park, D. (2021). Youtube as a source of patient information regarding exercises and compensated maneuvers for dysphagia. *Healthcare (Switzerland)*, *9*(8). https://doi.org/10.3390/HEALTHCARE9081084/S1

Charnock, D. (1998). *The Discern Handbook* (Radcliffe Medical Press (ed.)).

Charnock, D., Shepperd, S., Needham, G., & Gann, R. (1999). DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology & Community Health*, *53*(2), 105–111. https://doi.org/10.1136/jech.53.2.105

Chen, P., Mao, L., Nassis, G. P., Harmer, P., Ainsworth, B. E., & Li, F. (2020). Wuhan coronavirus (2019-nCoV): The need to maintain regular physical activity while taking precautions. *Journal of Sport and Health Science*, *9*(2), 103–104. https://doi.org/10.1016/j.jshs.2020.02.001

Chu, C. S., Lee, N. P., Adeoye, J., Thomson, P., & Choi, S. W. (2020). Machine learning and treatment outcome prediction for oral cancer. *Journal of Oral Pathology and Medicine*, *49*(10), 977–985. https://doi.org/10.1111/jop.13089

Croux, C., Dehon, C., Croux, C., & Dehon, C. (2010). Influence functions of the Spearman and Kendall correlation measures. *Statistical Methods & Applications 2010 19:4*, *19*(4), 497–515. https://doi.org/10.1007/S10260-010-0142-Z

Crutchfield, C. R., Frank, J. S., Anderson, M. J., Trofa, D. P., & Lynch, T. S. (2021). A Systematic Assessment of YouTube Content on Femoroacetabular Impingement: An Updated Review. *Orthopaedic Journal of Sports Medicine*, *9*(6). https://doi.org/10.1177/23259671211016340

Culha, Y., Seyhan Ak, E., Merder, E., Ariman, A., & Culha, M. G. (2021). Analysis of the YouTube videos on pelvic floor muscle exercise training in terms of their reliability and quality. *International Urology and Nephrology*, *53*(1), 1–6. https://doi.org/10.1007/s11255-020-02620-w

de Lira, G. H. S., Fornari, A., Cardoso, L. F., Aranchipe, M., Kretiska, C., & Rhoden, E. L. (2019). Effects of perioperative pelvic floor muscle training on early recovery of urinary continence and erectile function in men undergoing radical prostatectomy: A randomized clinical trial. *International Braz J Urol*, *45*(6), 1196–1203. https://doi.org/10.1590/S1677-5538.IBJU.2019.0238

Di Bello, F., Collà Ruvolo, C., Cilio, S., La Rocca, R., Capece, M., Creta, M., Celentano, G., Califano, G., Morra, S., Iacovazzo, C., Coviello, A., Buonanno, P., Fusco, F., Imbimbo, C., Mirone, V., & Longo, N. (2022). Testicular cancer and YouTube: What do you expect from a social media platform? *International Journal of Urology : Official Journal of the Japanese Urological Association*, *29*(7), 685–691. https://doi.org/10.1111/IJU.14871

Díaz de León Castañeda, C., & Martínez Domínguez, M. (2021). Factors Related to Internet Adoption and Its Use to Seek Health Information in Mexico. *Health Communication*, *36*(13), 1–8. https://doi.org/10.1080/10410236.2020.1794552

Divine, G., Norton, H. J., Hunt, R., & Dienemann, J. (2013). A review of analysis and sample size calculation considerations for wilcoxon tests. *Anesthesia and Analgesia*, *117*(3), 699–710. https://doi.org/10.1213/ANE.0B013E31827F53D7

Drozd, B., Couvillon, E., & Suarez, A. (2018). Medical YouTube Videos and Methods of Evaluation: Literature Review. *JMIR Med Educ 2018;4(1):E3*

Https://Mededu.Jmir.Org/2018/1/E3, *4*(1), e8527. https://doi.org/10.2196/MEDEDU.8527

Dubey, D., Amritphale, A., Sawhney, A., Dubey, D., & Srivastav, N. (2014). Analysis of YouTube as a Source of Information for West Nile Virus Infection. *Clinical Medicine & Research*, *12*(3–4), 129–132. https://doi.org/10.3121/CMR.2013.1194

Duran, M. B., & Kizilkan, Y. (2021). Quality analysis of testicular cancer videos on YouTube. *Andrologia*, *53*(8). https://doi.org/10.1111/AND.14118

Ďuriš, V., Bartková, R., & Tirpáková, A. (2021). Principal component analysis and factor analysis for an atanassov if data set. *Mathematics*, *9*(17), 2067. https://doi.org/10.3390/math9172067

Dweekat, O. Y., & Lam, S. S. (2022). Cervical Cancer Diagnosis Using an Integrated System of Principal Component Analysis, Genetic Algorithm, and Multilayer Perceptron. *Healthcare (Switzerland)*, *10*(10). https://doi.org/10.3390/healthcare10102002

Eastment, H. T., & Krzanowski, W. J. (1982). Cross-validatory choice of the number of components from a principal component analysis. *Technometrics*, *24*(1), 73–77. https://doi.org/10.1080/00401706.1982.10487712

Efe, S. C., Kalkan, S., & Karagoz, A. (2021). Evaluation of pericardiocentesis videos on YouTube as a reliable source for on-line education in the COVID-19 period. *Kardiochirurgia i Torakochirurgia Polska = Polish Journal of Cardio-Thoracic Surgery*, *18*(1), 33–39. https://doi.org/10.5114/KITP.2021.105185

Enver, N., Doruk, C., Kara, H., Gürol, E., Incaz, S., & Mamadova, U. (2020). YouTube™ as an information source for larynx cancer: a systematic review of video content. *European Archives of Oto-Rhino-Laryngology : Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*, *277*(7), 2061–2069. https://doi.org/10.1007/S00405-020-05906-Y

Erdem, M. N., & Karaca, S. (2018). Evaluating the accuracy and quality of the information in kyphosis videos shared on youtube. *Spine*, *43*(22), E1334–E1339. https://doi.org/10.1097/BRS.0000000000002691

Esen, E., Aslan, M., Sonbahar, B. Ç., & Kerimoğlu, R. S. (2019). YouTube English videos as a source of information on breast self-examination. *Breast Cancer Research and Treatment*, *173*(3), 629–635. https://doi.org/10.1007/s10549-018-5044-z

Fahy, E., Hardikar, R., Fox, A., & Mackay, S. (2014). Quality of patient health information on the internet: Reviewing a complex and evolving landscape. *Australasian Medical Journal*, *7*(ue 1), 24–28. https://doi.org/10.4066/AMJ.2014.1900

Fisher, A. J., Howe, E., & Zong, Z. Y. (2022). Unsupervised clustering of autonomic temporal networks in clinically distressed and psychologically healthy individuals. *Behaviour Research and Therapy*, *154*. https://doi.org/10.1016/J.BRAT.2022.104105

Fode, M., Nolsøe, A. B., Jacobsen, F. M., Russo, G. I., Østergren, P. B., Jensen, C. F. S., Albersen, M., Capogrosso, P., & Sønksen, J. (2020). Quality of Information in YouTube Videos on Erectile Dysfunction. *Sexual Medicine*, *8*(3), 408–413. https://doi.org/10.1016/J.ESXM.2020.05.007

Gad, B., Shanmugasundaram, S., Kumar, A., & Shukla, P. (2022). Quality and Reliability of YouTube Videos on Uterine Fibroid Embolization. *Journal of the American College of Radiology : JACR*, *19*(7), 905–912. https://doi.org/10.1016/J.JACR.2022.03.014

Gautam, A. P., Maiya, A. G., & Vidyasagar, M. S. (2011). Effect of home-based exercise program on lymphedema and quality of life in female postmastectomy patients: Pre-post intervention study. *Journal of Rehabilitation Research & Development*, *48*(10), 1261–1268. https://doi.org/10.1682/JRRD.2010.05.0089

Gewers, F. L., Ferreira, G. R., De Arruda, H. F., Silva, F. N., Comin, C. H., Amancio, D. R., & Costa, L. D. F. (2021). Principal Component Analysis. *ACM Computing Surveys (CSUR)*, *54*(4). https://doi.org/10.1145/3447755

Gokcen, H. B., & Gumussuyu, G. (2019). A Quality Analysis of Disc Herniation Videos on YouTube. *World Neurosurgery*, *124*, e799–e804. https://doi.org/10.1016/j.wneu.2019.01.146

Goobie, G. C., Guler, S. A., Johannson, K. A., Fisher, J. H., & Ryerson, C. J. (2019). YouTube Videos as a Source of Misinformation on Idiopathic Pulmonary Fibrosis. *Annals of the American Thoracic Society*, *16*(5), 572–579. https://doi.org/10.1513/ANNALSATS.201809-644OC

Güloğlu, S., Özdemir, Y., Basim, P., & Tolu, S. (2022). YouTube English videos as a source of information on arm and shoulder exercise after breast cancer surgery. *European Journal of Cancer Care*. https://doi.org/10.1111/ECC.13685

Gulve, N. D., Tripathi, P. R., Dahivelkar, S. D., Gulve, M. N., Gulve, R. N., & Kolhe, S. J. (2022). Evaluation of YouTube Videos as a Source of Information About Oral Self-examination to Detect Oral Cancer and Precancerous Lesions. *Journal of International Society of*

Preventive & Community Dentistry, *12*(2), 226–234. https://doi.org/10.4103/JISPCD.JISPCD_277_21

Gun, C., Aldinc, H., Ugur, E. M., Cadircibasi, E. R., & Yaylaci, S. (2022). Should Emergency Physicians and Nurses Direct Their Patients to YouTube for Heparin Self-Injection Training? A Systematic Review of Social Media Videos. *Journal of Emergency Nursing*, *48*(4), 376–389. https://doi.org/10.1016/J.JEN.2022.03.007

Gunasekeran, D. V., Tham, Y. C., Ting, D. S. W., Tan, G. S. W., & Wong, T. Y. (2021). Digital health during COVID-19: lessons from operationalising new models of care in ophthalmology. *The Lancet Digital Health*, *3*(2), e124–e134. https://doi.org/10.1016/S2589-7500(20)30287-9

Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism: Clinical and Experimental*, *69*, S36–S40. https://doi.org/10.1016/j.metabol.2017.01.011

Hamzehei, R., Ansari, M., Rahmatizadeh, S., & Valizadeh-Haghi, S. (2018). Websites as a tool for public health education: determining the trustworthiness of health websites on Ebola disease. *Online Journal of Public Health Informatics*, *10*(3). https://doi.org/10.5210/OJPHI.V10I3.9544

Hartley, J. (2012). Digital Futures for Cultural and Media Studies. In *Digital Futures for Cultural and Media Studies*. https://doi.org/10.1002/9781118106723

Hasan, H., & Tahir, N. M. (2010). Feature selection of breast cancer based on Principal Component Analysis. *Proceedings - CSPA 2010: 2010 6th International Colloquium on Signal Processing and Its Applications*, 242–245. https://doi.org/10.1109/CSPA.2010.5545298

Hodges, P. W., Stafford, R. E., Hall, L., Neumann, P., Morrison, S., Frawley, H., Doorbar-Baptist, S., Nahon, I., Crow, J., Thompson, J., & Cameron, A. P. (2020). Reconsideration of pelvic floor muscle training to prevent and treat incontinence after radical prostatectomy. *Urologic Oncology: Seminars and Original Investigations*, *38*(5), 354–371. https://doi.org/10.1016/J.UROLONC.2019.12.007

HONcode: E-Guide for Health Consumers. (2020). https://www.hon.ch/HONcode/Pro/HealthEvaluationTool.html

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

International Telecommunication Union ITU. (2022). *Global Connectivity Report 2022 - ITU Publication*. https://www.itu.int/hub/publication/d-ind-global-01-2022/

Internet World Stats. (2022). *World Internet Users Statistics and 2022 World Population Stats*. Www.Internetworldstats.Com. https://www.internetworldstats.com/stats.htm

Ivosev, G., Burton, L., & Bonner, R. (2008). Dimensionality reduction and visualization in principal component analysis. *Analytical Chemistry*, *80*(13), 4933–4944. https://doi.org/10.1021/AC800110W/ASSET/IMAGES/LARGE/AC-2008-00110W_0008.JPEG

Jackson, J. E. (1991). *A Use's Guide to Principal Components*. https://doi.org/10.1002/0471725331

Jain, A., & Dubes, R. (1988). *Algorithms for clustering data*. https://dl.acm.org/doi/abs/10.5555/42779

Jessen, M., Lorenz, C., Boehm, E., Hertling, S., Hinz, M., Imiolczyk, J.-P., Pelz, C., Ameziane, Y., & Lappen, S. (2022). Patient education on subacromial impingement syndrome : Reliability and educational quality of content available on Google and YouTube. *Orthopadie (Heidelberg, Germany)*. https://doi.org/10.1007/S00132-022-04294-X

Jha, S., & Topol, E. J. (2016). Adapting to Artificial Intelligence: Radiologists and Pathologists as Information Specialists. *JAMA*, *316*(22), 2353–2354. https://doi.org/10.1001/JAMA.2016.17438

Johnson, K. W., Torres Soto, J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., Ashley, E., & Dudley, J. T. (2018). Artificial Intelligence in Cardiology. *Journal of the American College of Cardiology*, *71*(23), 2668–2679. https://doi.org/10.1016/J.JACC.2018.03.521

Jolliffe, I. (2002). *Principal component analysis for special types of data*. https://link.springer.com/content/pdf/10.1007/0-387-22440-8_13.pdf

Jolliffe, I. (2006a). Graphical Representation of Data Using Principal Components. In *Principal Component Analysis* (pp. 78–110). Springer, New York, NY. https://doi.org/10.1007/0-387-22440-8_5

Jolliffe, I. (2006b). Principal Component Analysis for Special Types of Data. In *Principal Component Analysis* (pp. 338–372). Springer, New York, NY. https://doi.org/10.1007/0-387-22440-8_13

Kadakia, S., Stratton, C., Wu, Y., Feliciano, J., & Tuakli-Wosornu, Y. A. (2022). The Accessibility of YouTube Fitness Videos for Individuals Who Are Disabled Before and During the COVID-19 Pandemic: Preliminary Application of a Text Analytics Approach. *JMIR Formative Research*, *6*(2). https://doi.org/10.2196/34176

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401–415. https://doi.org/10.1007/BF02291817

Kaiser, H. F. (1982). Educational and Psychological Measurement. *Serials Review*, *8*(3), 83–85. http://www.tandfonline.com/doi/abs/10.1080/00987913.1982.10763392

Kambhatla, N., & Leen, T. K. (1997). Dimension Reduction by Local Principal Component Analysis. *Neural Computation*, *9*(7), 1493–1516. https://doi.org/10.1162/neco.1997.9.7.1493

Kanyongo, G. Y. (2005). Determining The Correct Number Of Components To Extract From A Principal Components Analysis: A Monte Carlo Study Of The Accuracy Of The Scree Plot. *Journal of Modern Applied Statistical Methods*, *4*(1), 13. https://doi.org/10.22237/jmasm/1114906380

Karakoyun, A., & Yildirim, A. (2021). YouTube videos as a source of information concerning Behçet's disease: a reliability and quality analysis. *Rheumatology International*, *41*(12), 2117–2123. https://doi.org/10.1007/s00296-021-05009-9

Kataria, A., & Singh, M. D. (2008). A Review of Data Classification Using K-Nearest Neighbour Algorithm. In *International Journal of Emerging Technology and Advanced Engineering* (Vol. 3, Issue 6). https://www.researchgate.net/publication/353306410

Katz, M., & Nandi, N. (2021). Social Media and Medical Education in the Context of the COVID-19 Pandemic: Scoping Review. *JMIR Medical Education*, *7*(2). https://doi.org/10.2196/25892

Kegel, A. H. (1948). Progressive resistance exercise in the functional restoration of the perineal muscles. *American Journal of Obstetrics and Gynecology*, *56*(2), 238–248. https://doi.org/10.1016/0002-9378(48)90266-X

Kılınç, D. D., & Sayar, G. (2019). Assessment of reliability of YouTube videos on orthodontics. *Turkish Journal of Orthodontics*, *32*(3), 145–150. https://doi.org/10.5152/TurkJOrthod.2019.18064

Kocyigit, B. F., Nacitarhan, V., Koca, T. T., & Berk, E. (2019). YouTube as a source of patient information for ankylosing spondylitis exercises. *Clinical Rheumatology*, *38*(6), 1747–1751. https://doi.org/10.1007/s10067-018-04413-0

Kodonas, K., & Fardi, A. (2021). YouTube as a source of information about pulpotomy and pulp capping: a cross sectional reliability analysis. *Restorative Dentistry & Endodontics*, *46*(3). https://doi.org/10.5395/RDE.2021.46.E40

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation

Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kraus, S., Schiavone, F., Pluzhnikova, A., & Invernizzi, A. C. (2021). Digital transformation in healthcare: Analyzing the current state-of-research. *Journal of Business Research*, *123*, 557–567. https://doi.org/10.1016/J.JBUSRES.2020.10.030

Kunze, K. N., Cohn, M. R., Wakefield, C., Hamati, F., LaPrade, R. F., Forsythe, B., Yanke, A. B., & Chahla, J. (2019). YouTube as a Source of Information About the Posterior Cruciate Ligament: A Content-Quality and Reliability Analysis. *Arthroscopy, Sports Medicine, and Rehabilitation*, *1*(2), e109–e114. https://doi.org/10.1016/j.asmr.2019.09.003

Kuru, T., & Erken, H. Y. (2020). Evaluation of the Quality and Reliability of YouTube Videos on Rotator Cuff Tears. *Cureus*, *12*(2). https://doi.org/10.7759/cureus.6852

Laacke, S., Mueller, R., Schomerus, G., & Salloch, S. (2021). Artificial Intelligence, Social Media and Depression. A New Concept of Health-Related Digital Autonomy. *The American Journal of Bioethics*, *21*(7), 4–20. https://doi.org/10.1080/15265161.2020.1863515

Lang, J. J., Giffen, Z., Hong, S., Demeter, J., El-Zawahry, A., Sindhwani, P., & Ekwenna, O. (2022). Assessing Vasectomy-Related Information on YouTube: An Analysis of the Quality, Understandability, and Actionability of Information. *American Journal of Men's Health*, *16*(2). https://doi.org/10.1177/15579883221094716

Langford, A., & Loeb, S. (2019). Perceived Patient-Provider Communication Quality and Sociodemographic Factors Associated With Watching Health-Related Videos on YouTube: A Cross-Sectional Analysis. *J Med Internet Res 2019;21(5):E13512 Https://Www.Jmir.Org/2019/5/E13512*, *21*(5), e13512. https://doi.org/10.2196/13512

Lau, K. Y. Y., Ng, K. S., Kwok, K. W., Tsia, K. K. M., Sin, C. F., Lam, C. W., & Vardhanabhuti, V. (2022). An Unsupervised Machine Learning Clustering and Prediction of Differential Clinical Phenotypes of COVID-19 Patients Based on Blood Tests—A Hong Kong Population Study. *Frontiers in Medicine*, *8*. https://doi.org/10.3389/FMED.2021.764934/PDF

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature 2015 521:7553*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Ledesma, R. D., Valero-Mora, P., & Macbeth, G. (2015). The Scree Test and the Number of Factors: a Dynamic Graphics Approach. *The Spanish Journal of Psychology*, *18*, E11.

https://doi.org/10.1017/SJP.2015.13

Lei, Y.-M., Yin, M., Yu, M.-H., Yu, J., Zeng, S.-E., Lv, W.-Z., Li, J., Ye, H.-R., Cui, X.-W., & Dietrich, C. F. (2021). Artificial Intelligence in Medical Imaging of the Breast. *Frontiers in Oncology*, *11*. https://doi.org/10.3389/fonc.2021.600557

Lewis, S. P., Heath, N. L., Sornberger, M. J., & Arbuthnott, A. E. (2012). Helpful or harmful? An examination of viewers' responses to nonsuicidal self-injury videos on youtube. *Journal of Adolescent Health*, *51*(4), 380–385. https://doi.org/10.1016/j.jadohealth.2012.01.013

Li, L., Cheng, W. Y., Glicksberg, B. S., Gottesman, O., Tamler, R., Chen, R., Bottinger, E. P., & Dudley, J. T. (2015). Identification of type 2 diabetes subgroups through topological analysis of patient similarity. *Science Translational Medicine*, *7*(311). https://doi.org/10.1126/SCITRANSLMED.AAA9364/SUPPL_FILE/7-311RA174_SM.PDF

Li, M., Shi, J., & Chen, Y. (2022). Identifying Influences in Patient Decision-making Processes in Online Health Communities: Data Science Approach. *J Med Internet Res 2022;24(8):E30634 Https://Www.Jmir.Org/2022/8/E30634*, *24*(8), e30634. https://doi.org/10.2196/30634

Loeb, S., Reines, K., Abu-Salha, Y., French, W., Butaney, M., Macaluso, J. N., Steinberg, G. D., Walter, D., Byrne, N., la Garza, D. de, & Smith, A. B. (2021). Quality of Bladder Cancer Information on YouTube. *European Urology*, *79*(1), 56–59. https://doi.org/10.1016/j.eururo.2020.09.014

Loeb, S., Sengupta, S., Butaney, M., Macaluso, J. N., Czarniecki, S. W., Robbins, R., Braithwaite, R. S., Gao, L., Byrne, N., Walter, D., & Langford, A. (2019). Dissemination of Misinformative and Biased Information about Prostate Cancer on YouTube. *European Urology*, *75*(4), 564–567. https://doi.org/10.1016/j.eururo.2018.10.056

Lu, S., Christie, G. A., Nguyen, T. T., Freeman, J. D., & Hsu, E. B. (2022). Applications of Artificial Intelligence and Machine Learning in Disasters and Public Health Emergencies. *Disaster Medicine and Public Health Preparedness*, *16*(4), 1674–1681. https://doi.org/10.1017/dmp.2021.125

Luo, Y., Ahmad, F. S., & Shah, S. J. (2017). Tensor Factorization for Precision Medicine in Heart Failure with Preserved Ejection Fraction. *Journal of Cardiovascular Translational Research 2017 10:3*, *10*(3), 305–312. https://doi.org/10.1007/S12265-016-9727-8

Madathil, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S., & Gramopadhye, A. K. (2015). Healthcare information on YouTube: A systematic review. *Health Informatics Journal*, *21*(3), 173–194. https://doi.org/10.1177/1460458213512220

Mahmoudi, M. R., Heydari, M. H., Qasem, S. N., Mosavi, A., & Band, S. S. (2021). Principal component analysis to study the relations between the spread rates of COVID-19 in high risks countries. *Alexandria Engineering Journal*, *60*(1), 457–464. https://doi.org/10.1016/J.AEJ.2020.09.013

Mak, K. K., & Pichika, M. R. (2019). Artificial intelligence in drug development: present status and future prospects. *Drug Discovery Today*, *24*(3), 773–780. https://doi.org/10.1016/J.DRUDIS.2018.11.014

Mamlin, B. W., & Tierney, W. M. (2016). The Promise of Information and Communication Technology in Healthcare: Extracting Value from the Chaos. *American Journal of the Medical Sciences*, *351*(1), 59–68. https://doi.org/10.1016/j.amjms.2015.10.015

Marr, D. (1977). Artificial intelligence-A personal view. *Artificial Intelligence*, *9*(1), 37–48. https://doi.org/10.1016/0004-3702(77)90013-3

Marwah, H. K., Carlson, K., Rosseau, N. A., Chretien, K. C., Kind, T., & Jackson, H. T. (2021). Videos, Views, and Vaccines: Evaluating the Quality of COVID-19 Communications on YouTube. *Disaster Medicine and Public Health Preparedness*. https://doi.org/10.1017/DMP.2021.284

Matias, T., Dominski, F. H., & Marks, D. F. (2020). Human needs in COVID-19 isolation. In *Journal of Health Psychology* (Vol. 25, Issue 7, pp. 871–882). https://doi.org/10.1177/1359105320925149

McDonough, D. J., Helgeson, M. A., Liu, W., & Gao, Z. (2022). Effects of a remote, YouTube-delivered exercise intervention on young adults' physical activity, sedentary behavior, and sleep during the COVID-19 pandemic: Randomized controlled trial. *Journal of Sport and Health Science*, *11*(2), 145–156. https://doi.org/10.1016/J.JSHS.2021.07.009

McMahon, K. M., Schwartz, J., Nilles-Melchert, T., Ray, K., Eaton, V., & Chakkalakal, D. (2022). YouTube and the Achilles Tendon: An Analysis of Internet Information Reliability and Content Quality. *Cureus*, *14*(4). https://doi.org/10.7759/CUREUS.23984

Memioglu, T., & Ozyasar, M. (2022). Analysis of YouTube videos as a source of information for myocarditis during the COVID-19 pandemic. *Clinical Research in Cardiology : Official Journal of the German Cardiac Society*. https://doi.org/10.1007/S00392-022-02026-X

Meskó, B., & Görög, M. (2020). A short guide for medical professionals in the era of artificial intelligence. *Npj Digital Medicine 2020 3:1*, *3*(1), 1–8. https://doi.org/10.1038/s41746-020-00333-z

Meteran, H., Høj, S., Sigsgaard, T., Diers, C. S., Remvig, C., & Meteran, H. (2022). The usefulness of YouTube videos on lung cancer. *Journal of Public Health (Oxford, England)*. https://doi.org/10.1093/PUBMED/FDAC092

Milios, J. E., Ackland, T. R., & Green, D. J. (2019). Pelvic floor muscle training in radical prostatectomy: a randomized controlled trial of the impacts on pelvic floor muscle function and urinary incontinence. *BMC Urology*, *19*(166), 1–10. https://doi.org/10.1186/s12894-019-0546-5

Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Scientific Reports*, *6*. https://doi.org/10.1038/SREP26094

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. R. I. S. M. A. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*, *6*, 1000097. https://doi.org/10.1371/journal.pmed.1000097.

Morra, S., Collà Ruvolo, C., Napolitano, L., La Rocca, R., Celentano, G., Califano, G., Creta, M., Capece, M., Turco, C., Cilio, S., Melchionna, A., Gerundo, G., Trama, F., Mangiapia, F., Fusco, F., Mirone, V., & Longo, N. (2022). YouTubeTM as a source of information on bladder pain syndrome: A contemporary analysis. *Neurourology and Urodynamics*, *41*(1), 237–245. https://doi.org/10.1002/NAU.24802

Nangia, A. K., Selvi, I., Baydilli, N., & Akinsal, E. C. (2020). Can YouTube English Videos Be Recommended as an Accurate Source for Learning About Testicular Self-examination? COMMENT. *UROLOGY*, *145*, 188. https://doi.org/10.1016/j.urology.2020.06.085

Narayanan, A., Keedwell, E. C., & Olsson, B. (2002). Artificial intelligence techniques for bioinformatics. In *Applied bioinformatics* (Vol. 1, Issue 4, pp. 191–222). http://www.binf.gmu.edu/vaisman/binf630/ab02-narayanan-ai.pdf

Navarro, O., Escrivá, M., Faubel, R., & Traver, V. (2021). Empowering Patients Living With Chronic Conditions Using Video as an Educational Tool: Scoping Review. *Journal of Medical Internet Research*, *23*(7). https://doi.org/10.2196/26427

O'Callaghan, M. E., Raymond, E., Campbell, J., Vincent, A. D., Beckmann, K., Roder, D., Evans, S., McNeil, J., Millar, J., Zalcberg, J., Borg, M., & Moretti, K. (2017). Tools for predicting patient-reported outcomes in prostate cancer patients undergoing radical prostatectomy: A systematic review of prognostic accuracy and validity. *Prostate Cancer and Prostatic Diseases*, *20*(4), 378–388. https://doi.org/10.1038/pcan.2017.28

Omnicore Agency. (2022). *YouTube by the Numbers (2022): Stats, Demographics & Fun Facts*. Omnicoreagency.Com. https://www.omnicoreagency.com/youtube-statistics/

Onder, M. E., & Zengin, O. (2021). Quality of healthcare information on YouTube: psoriatic arthritis. *Zeitschrift Fur Rheumatologie*. https://doi.org/10.1007/S00393-021-01069-1

Osman, W., Mohamed, F., Elhassan, M., & Shoufan, A. (2022). Is YouTube a reliable source of health-related information? A systematic review. *BMC Medical Education*, *22*(1). https://doi.org/10.1186/S12909-022-03446-Z

Oydanich, M., Shah, Y., Shah, K., & Khouri, A. S. (2022). An Analysis of the Quality, Reliability, and Popularity of YouTube Videos on Glaucoma. *Ophthalmology. Glaucoma*, *5*(3), 306–312. https://doi.org/10.1016/J.OGLA.2021.10.002

Ozduran, E., & Büyükçoban, S. (2022). A content analysis of the reliability and quality of Youtube videos as a source of information on health-related post-COVID pain. *PeerJ*, *10*, e14089. https://doi.org/10.7717/peerj.14089

Paige, S. R., Stellefson, M., Krieger, J. L., Anderson-Lewis, C., Cheong, J. W., & Stopka, C. (2018). Proposing a Transactional Model of eHealth Literacy: Concept Analysis. *J Med Internet Res 2018;20(10):E10175 Https://Www.Jmir.Org/2018/10/E10175*, *20*(10), e10175. https://doi.org/10.2196/10175

Pamukcu, M., & Izci Duran, T. (2021). Are YouTube videos enough to learn anakinra self-injection? *Rheumatology International*, *41*(12), 2125–2131. https://doi.org/10.1007/S00296-021-04999-W

Pan, L. H., Lin, M. H., Pang, S. T., Wang, J., & Shih, W. M. (2019). Improvement of Urinary Incontinence, Life Impact, and Depression and Anxiety With Modified Pelvic Floor Muscle Training After Radical Prostatectomy. *American Journal of Men's Health*, *13*(3). https://doi.org/10.1177/1557988319851618

Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. *Journal of Global Health*, *8*(2), 20303. https://doi.org/10.7189/JOGH.08.020303

Park, D. I. (2021). Development and Validation of a Knowledge, Attitudes and Practices Questionnaire on COVID-19 (KAP COVID-19). *International Journal of Environmental Research and Public Health*, *18*(14). https://doi.org/10.3390/IJERPH18147493

Parker, K., Uddin, R., Ridgers, N. D., Brown, H., Veitch, J., Salmon, J., Timperio, A., Sahlqvist, S., Cassar, S., Toffoletti, K., Maddison, R., & Arundell, L. (2021). The Use of Digital Platforms for Adults' and Adolescents' Physical Activity During the COVID-19 Pandemic (Our Life at Home): Survey Study. *Journal of Medical Internet Research*, *23*(2). https://doi.org/10.2196/23389

Passos, K.-K., Leonel, A.-C., Bonan, P.-R., Castro, J.-F., Pontual, M.-L., Ramos-Perez, F.-M., & Perez, D.-E. (2020). Quality of information about oral cancer in Brazilian Portuguese available on Google, Youtube, and Instagram. *Medicina Oral, Patologia Oral y Cirugia Bucal*, *25*(3), e346–e352. https://doi.org/10.4317/medoral.23374

Patel, R. D., Abramowitz, C., Shamsian, E., Okhawere, K. E., Deluxe, A., Ayo-Farai, O., Korn, T. G., Meilika, K. N., & Badani, K. K. (2022). Is YouTube a good resource for patients to better understand kidney cancer? *Urologic Oncology*, *40*(6), 275.e19-275.e27. https://doi.org/10.1016/J.UROLONC.2022.02.023

Payedimarri, A. B., Concina, D., Portinale, L., Canonico, M., Seys, D., Vanhaecht, K., & Panella, M. (2021). Prediction Models for Public Health Containment Measures on COVID-19 Using Artificial Intelligence and Machine Learning: A Systematic Review. *International Journal of Environmental Research and Public Health*, *18*(9), 4499. https://doi.org/10.3390/ijerph18094499

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572. https://doi.org/10.1080/14786440109462720

Pechenizkiy, M., Tsymbal, A., & Puuronen, S. (2004). PCA-based feature transformation for classification: Issues in medical diagnostics. *Proceedings of the IEEE Symposium on Computer-Based Medical Systems*, *17*, 535–540. https://doi.org/10.1109/CBMS.2004.1311770

Piskin, E., Çolakoğlu, M. K., Bal, A., Oter, V., & Bostanci, E. B. (2021). Evaluation of the Quality, Educational Value and Utility of the Videos on YouTube for Laparoscopic Low Anterior Resection. *American Surgeon*, 31348211011102. https://doi.org/10.1177/00031348211011102

Prota, C., Gomes, C. M., Ribeiro, L. H. S., De Bessa, J., Nakano, E., Dall'oglio, M., Bruschini, H., & Srougi, M. (2012). Early postoperative pelvic-floor biofeedback improves erectile function in men undergoing radical prostatectomy: A prospective, randomized, controlled trial. *International Journal of Impotence Research*, *24*(5), 174–178. https://doi.org/10.1038/ijir.2012.11

Puth, M. T., Neuhäuser, M., & Ruxton, G. D. (2015). Effective use of Spearman's and Kendall's correlation coefficients for association between two measured traits. *Animal Behaviour*, *102*, 77–84. https://doi.org/10.1016/J.ANBEHAV.2015.01.010

Qu, Y., Ostrouchov, G., Samatova, N., & Geist, A. (2002). Principal Component Analysis for Dimension Reduction in Massive Distributed Data Sets. *Workshop on High Performance Data Mining at the Second SIAM International Conference on Data Mining*, *April*, 4–9. https://www.researchgate.net/publication/232063041

Radadiya, D., Gonzalez-Estrada, A., Lira-Vera, J. E., Lizarrga-Torres, K., Mahapatra, S. S., Murguia-Fuentes, R., Niezen, S., Mohan, K., Reddy, C., & Devani, K. (2020). Colonoscopy videos on YouTube: Are they a good source of patient education? *Endoscopy International Open*, *8*(5), E598–E606. https://doi.org/10.1055/a-1122-8566

Rashidi, H. H., Tran, N. K., Betts, E. V., Howell, L. P., & Green, R. (2019). Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. *Academic Pathology*, *6*, 2374289519873088. https://doi.org/10.1177/2374289519873088

Reigle, B. S., & Zhang, B. (2018). Women's Rehabilitation Experiences Following Breast Cancer Surgery. *Rehabilitation Nursing*, *43*(4), 195–200. https://doi.org/10.1097/rnj.0000000000000168

Richmond, H., Lait, C., Srikesavan, C., Williamson, E., Moser, J., Newman, M., Betteley, L., Fordham, B., Rees, S., Lamb, S. E., & Bruce, J. (2018). Development of an exercise intervention for the prevention of musculoskeletal shoulder problems after breast cancer treatment: the prevention of shoulder problems trial (UK PROSPER. *BMC Health Services Research*, *18*(1), 1–12. https://doi.org/10.1186/s12913-018-3280-x

Rodriguez-Rodriguez, A. M., Blanco-Diaz, M., de la Fuente-Costa, M., Hernandez-Sanchez, S., Escobio-Prieto, I., & Casaña, J. (2022). Review of the Quality of YouTube Videos Recommending Exercises for the COVID-19 Lockdown. In *International Journal of Environmental Research and Public Health* (Vol. 19, Issue 13, p. 8016). Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/ijerph19138016

Rodriguez-Rodriguez, A. M., Blanco-Diaz, M., Lopez-Diaz, P., de la Fuente-Costa, M., Duenas, L., Prieto, I. E., Calatayud, J., & Casana-Granell, J. (2021). Quality Analysis of YouTube Videos Presenting Shoulder Exercises after Breast Cancer Surgery. *Breast Care*. https://doi.org/10.1159/000518265

Rodriguez-Rodriguez, A. M., Blanco-Diaz, M., Lopez-Diaz, P., de la Fuente-Costa, M., Sousa-Fraguas, M. C., Escobio-Prieto, I., & Casaña, J. (2021). Quality analysis of youtube videos presenting pelvic floor exercises after prostatectomy surgery. *Journal of Personalized Medicine*, *11*(9), 920. https://doi.org/10.3390/jpm11090920

Rodríguez, A. H., Ruiz-Botella, M., Martín-Loeches, I., Jimenez Herrera, M., Solé-Violan, J., Gómez, J., Bodí, M., Trefler, S., Papiol, E., Díaz, E., Suberviola, B., Vallverdu, M., Mayor-Vázquez, E., Albaya Moreno, A., Canabal Berlanga, A., Sánchez, M., del Valle Ortíz, M., Ballesteros, J. C., Martín Iglesias, L., … Ribas, A. M. (2021). Deploying unsupervised clustering analysis to derive clinical phenotypes and risk factors associated with mortality risk in 2022 critically ill patients with COVID-19 in Spain. *Critical Care (London, England)*, *25*(1). https://doi.org/10.1186/S13054-021-03487-8

Roohi, A., Faust, K., Djuric, U., & Diamandis, P. (2020). Unsupervised Machine Learning in Pathology: The Next Frontier. In *Surgical Pathology Clinics* (Vol. 13, Issue 2, pp. 349–358). W.B. Saunders. https://doi.org/10.1016/j.path.2020.01.002

Ryu, J., Kim, H., Kim, R. M., Kim, S., Jo, J., Lee, S., Nam, K. T., Joo, Y. C., Yi, G. C., Lee, J., & Kim, M. (2021). Dimensionality reduction and unsupervised clustering for EELS-SI. *Ultramicroscopy*, *231*. https://doi.org/10.1016/J.ULTRAMIC.2021.113314

Saerens, M., Fouss, F., Yen, L., & Dupont, P. (2004). The principal components analysis of a graph, and its relationships to spectral clustering. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, *3201*, 371–383. https://doi.org/10.1007/978-3-540-30115-8_35/COVER

Salih Hasan, B. M., & Abdulazeez, A. M. (2021). A Review of Principal Component Analysis Algorithm for Dimensionality Reduction. *Journal of Soft Computing and Data Mining*, *02*(01), 20–30. https://doi.org/10.30880/jscdm.2021.02.01.003

Salman, M. Y., & Bayar, G. (2021). Evaluation of quality and reliability of YouTube videos on female urinary incontinence. *Journal of Gynecology Obstetrics and Human Reproduction*, *50*(10). https://doi.org/10.1016/J.JOGOH.2021.102200

Şan, H. (2022). Use of YouTube as an Information Source for Radioactive Iodine Therapy: Do YouTube Videos Have High Quality? *Molecular Imaging and Radionuclide Therapy*, *31*(1), 42–48. https://doi.org/10.4274/mirt.galenos.2021.85530

Schwalbe, N., & Wahl, B. (2020). Artificial intelligence and the future of global health. *The Lancet*, *395*(10236), 1579–1586. https://doi.org/10.1016/S0140-6736(20)30226-9

Shah, S. G. S., Farrow, A., Purcell, L. N., & Charles, A. G. (2020). "World Health Organization declares global emergency: A review of the 2019 novel Coronavirus (COVID-19)." *International Journal of Surgery*, *76*(January), 128–129. https://doi.org/10.1016/j.ijsu.2020.03.001

Shahid, N., Kalofolias, V., Bresson, X., Bronstein, M., & Vandergheynst, P. (2015). Robust principal component analysis on graphs. *Proceedings of the IEEE International Conference on Computer Vision*, *2015 Inter*, 2812–2820. https://doi.org/10.1109/ICCV.2015.322

Shameer, K., Badgeley, M. A., Miotto, R., Glicksberg, B. S., Morgan, J. W., & Dudley, J. T. (2017). Translational bioinformatics in the era of real-time biomedical, health care and wellness data streams. *Briefings in Bioinformatics*, *18*(1), 105–124. https://doi.org/10.1093/BIB/BBV118

Shungu, N., Haley, S. P., Berini, C. R., Foster, D., & Diaz, V. A. (2021). Quality of YouTube videos on prostate cancer screening for black men. *Journal of the American Board of Family Medicine*, *34*(4), 724–731. https://doi.org/10.3122/JABFM.2021.04.200632

Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2021). Cancer Statistics, 2021. *CA: A Cancer Journal for Clinicians*, *71*(1), 7–33. https://doi.org/10.3322/caac.21654

Silva, R., Oliveira, D., Santos, D. P., Santos, L. F. D., Wilson, R. E., & Bedo, M. (2020). Criteria for choosing the number of dimensions in a principal component analysis: An empirical assessment. *Anais Do Simpósio Brasileiro de Banco de Dados (SBBD)*, 145–150. https://doi.org/10.5753/SBBD.2020.13632

Singh, J. A. (2019). Artificial Intelligence and global health: opportunities and challenges. *Emerging Topics in Life Sciences*, *3*(6), 741–746. https://doi.org/10.1042/ETLS20190106

Soto, J. T., Hershman, S. G., & Ashley, E. A. (2021). Combining digital data and artificial intelligence for cardiovascular health. *Cardiovascular Research*, *117*(9), e116–e117. https://doi.org/10.1093/cvr/cvab211

Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, *15*(1), 72. https://doi.org/10.2307/1412159

Statista. (2022). *YouTube: Statistics & Data*. https://www.statista.com/topics/2019/youtube/#dossierKeyfigures

Steeb, T., Reinhardt, L., Harlaß, M., Heppt, M. V., Meier, F., & Berking, C. (2022). Assessment of the Quality, Understandability, and Reliability of YouTube Videos as a Source of

Information on Basal Cell Carcinoma: Web-Based Analysis. *JMIR Cancer*, *8*(1). https://doi.org/10.2196/29581

Steinhubl, S. R., Muse, E. D., & Topol, E. J. (2015). The emerging field of mobile health. In *Science Translational Medicine* (Vol. 7, Issue 283, pp. 283–286). https://doi.org/10.1126/scitranslmed.aaa3487

Sui, W., Sui, A., & Rhodes, R. E. (2022). What to watch: Practical considerations and strategies for using YouTube for research. *Digital Health*, *8*. https://doi.org/10.1177/20552076221123707

Tajudeen, F. P., Bahar, N., Maw Pin, T., & Saedon, N. I. (2021). Mobile Technologies and Healthy Ageing: A Bibliometric Analysis on Publication Trends and Knowledge Structure of mHealth Research for Older Adults. *Https://Doi.Org/10.1080/10447318.2021.1926115*, *38*(2), 118–130. https://doi.org/10.1080/10447318.2021.1926115

The Global Cancer Observatory. (2020). *All cancers*. https://gco.iarc.fr/today

Thiébaut, R., & Thiessard, F. (2018). Artificial Intelligence in Public Health and Epidemiology. *Yearbook of Medical Informatics*, *27*(1), 207–210. https://doi.org/10.1055/S-0038-1667082

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vu, E., Pratsinis, M., Plasswilm, L., Schmid, H.-P., Panje, C., & Betschart, P. (2021). Radiotherapy or Surgery? Comparative, Qualitative Assessment of Online Patient Education Materials on Prostate Cancer. *Current Oncology (Toronto, Ont.)*, *28*(5), 3420–3429. https://doi.org/10.3390/curroncol28050296

Wang, J. X., Somani, S., Chen, J. H., Murray, S., & Sarkar, U. (2021). Health Equity in Artificial Intelligence and Primary Care Research: Protocol for a Scoping Review. *JMIR Research Protocols*, *10*(9), e27799. https://doi.org/10.2196/27799

Warren, C. J., Sawhney, R., Shah, T., Behbahani, S., & Sadeghi-Nejad, H. (2021). YouTube and Men's Health: A Review of the Current Literature. *Sexual Medicine Reviews*, *9*(2), 280–288. https://doi.org/10.1016/J.SXMR.2020.09.002

Warren, C. J., Wisener, J., Ward, B., Behbahani, S., Shah, T., Fano, A., Paskhover, B., & Sadeghi-Nejad, H. (2021). YouTube as a Patient Education Resource for Male Hypogonadism and Testosterone Therapy. *Sexual Medicine*, *9*(2). https://doi.org/10.1016/J.ESXM.2021.100324

Weber Shandwick AND KRC Research. (2018). *The great American search for healthcare information.* https://www.webershandwick.com/wp-content/uploads/2018/11/Healthcare-Info-Search-Report.pdf

Wilkens, F. M., Ganter, C., Kriegsmann, K., Wilkens, H., Kahn, N., Goobie, G. C., Ryerson, C. J., & Kreuter, M. (2022). YouTube-videos for patient education in lymphangioleiomyomatosis? *Respiratory Research*, *23*(1). https://doi.org/10.1186/S12931-022-02022-9

World Health Organization. (2010). Recomendaciones Mundiales sobre Actividad Física para la Salud. *Geneva: WHO Library Cataloguing-in-Publication*, *Completo*, 1–58. https://apps.who.int/iris/bitstream/handle/10665/44441/9789243599977_spa.pdf?ua=1

World Health Organization. (2020a). *How to stay physically active during COVID-19 self-quarantine*. Worl Health Organization (WHO). https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/publications-and-technical-guidance/noncommunicable-diseases/stay-physically-active-during-self-quarantine

World Health Organization. (2020b). *Physical activity*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/physical-activity

World Health Organization. (2022). WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. In *World Health Organization* (pp. 1–5). https://covid19.who.int/

Worthen A. (2022). *American Physical Therapy Association*. Physical Therapy. https://doi.org/10.1093/ptj/45.5.404

Yalkin, O., Uzunoglu, M. Y., Iflazoglu, N., & Esen, E. (2022). Are YouTube Videos a Reliable Source for Obtaining Information About Colorectal Cancer Screening? *CUREUS JOURNAL OF MEDICAL SCIENCE*, *14*(1). https://doi.org/10.7759/cureus.21550

Yang, Y., Xu, L., Sun, L., Zhang, P., & Farid, S. S. (2022). Machine learning application in personalised lung cancer recurrence and survivability prediction. *Computational and Structural Biotechnology Journal*, *20*, 1811–1820. https://doi.org/10.1016/j.csbj.2022.03.035

Yeung, K. Y., & Ruzzo, W. L. (2001). Principal component analysis for clustering gene expression data. *Bioinformatics*, *17*(9), 763–774. https://doi.org/10.1093/BIOINFORMATICS/17.9.763

Yildiz, S., & Toros, S. Z. (2021). The Quality, Reliability, and Popularity of YouTube Education

Videos for Vestibular Rehabilitation: A Cross-sectional Study. *Otology & Neurotology : Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *42*(8), E1077–E1083. https://doi.org/10.1097/MAO.0000000000003197

Yoldemir, T. (2020). Artificial intelligence and women's health. *Climacteric : The Journal of the International Menopause Society*, *23*(1), 1–2. https://doi.org/10.1080/13697137.2019.1682804

Yu, H., Lu, W., Sun, Q., Shi, H., Wei, J., Wang, Z., Wang, X., & Xiong, N. (2022). Design and analysis of a robust breast cancer diagnostic system based on multimode MR images. *Mathematical Biosciences and Engineering*, *18*(4), 3578–3597. https://doi.org/10.3934/MBE.2021180

Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction. *Journal of Applied Science and Technology Trends*, *1*(2), 56–70. https://doi.org/10.38094/jastt1224

Zhang, Shihui, Sun, X., Liu, N., & Mi, J. (2022). Human Health Risk Prediction Method of Regional Atmospheric Environmental Pollution Sources Based on PMF and PCA Analysis under Artificial Intelligence Cloud Model. *International Journal of Analytical Chemistry*, *2022*. https://doi.org/10.1155/2022/7207020

Zhang, Shun, Fukunaga, T., Oka, S., Orita, H., Kaji, S., Yube, Y., Yamauchi, S., Kohira, Y., & Egawa, H. (2020). Concerns of quality, utility, and reliability of laparoscopic gastrectomy for gastric cancer in public video sharing platform. *Annals of Translational Medicine*, *8*(5), 196. https://doi.org/10.21037/atm.2020.01.78

Zhang, T., & Yang, B. (2016). Big Data Dimension Reduction Using PCA. *Proceedings - 2016 IEEE International Conference on Smart Cloud, SmartCloud 2016*, 152–157. https://doi.org/10.1109/SMARTCLOUD.2016.33

Zhang, X., Yang, Y., Shen, Y.-W., Zhang, K.-R., Ma, L.-T., Ding, C., Wang, B.-Y., Meng, Y., & Liu, H. (2022). Quality of online video resources concerning patient education for neck pain: A YouTube-based quality-control study. *Frontiers in Public Health*, *10*. https://doi.org/10.3389/FPUBH.2022.972348/PDF

Zhang, Z., & Castelló, A. (2017). Principal components analysis in clinical studies. *Annals of Translational Medicine*, *5*(17). https://doi.org/10.21037/atm.2017.07.12

Zhou, K., Wang, W., An, J., Li, M., Li, J., & Li, X. (2019). Effects of Progressive Upper Limb Exercises and Muscle Relaxation Training on Upper Limb Function and Health-Related Quality of Life Following Surgery in Women with Breast Cancer: A Clinical Randomized Controlled Trial. *Annals of Surgical Oncology*, *26*(7), 2156–2165. https://doi.org/10.1245/s10434-019-07305-y

Zhu, M., & Ghodsi, A. (2006). Automatic dimensionality selection from the scree plot via the use of profile likelihood. *Computational Statistics and Data Analysis*, *51*(2), 918–930. https://doi.org/10.1016/j.csda.2005.09.010

Zhuang, S., Li, L., Zang, Y., Li, G., & Wang, F. (2020). RRM2 elicits the metastatic potential of breast cancer cells by regulating cell invasion, migration and VEGF expression via the PI3K/AKT signaling. *Oncology Letters*, *19*(4), 3349–3355. https://doi.org/10.3892/ol.2020.11428

Zieff, G., Bates, L. C., Kerr, Z. Y., Moore, J. B., Hanson, E. D., Battaglini, C., & Stoner, L. (2021). Targeting sedentary behavior as a feasible health strategy during COVID-19. In *Translational Behavioral Medicine* (Vol. 11, Issue 3, pp. 826–831). https://doi.org/10.1093/tbm/ibaa101