





## REVISIÓN SISTEMÁTICA

# Big Data in Health Information Systems

## Big Data en los Sistemas de Información en Salud

Carlos Rafael Araujo Inastrilla<sup>1</sup>  

<sup>1</sup>Universidad de Ciencias Médicas de la Habana. Facultad de Tecnología de la Salud. La Habana, Cuba.

**Citar como:** Araujo Inastrilla CR. Big Data in Health Information Systems. *Seminars in Medical Writing and Education* 2022; 1:6. <https://doi.org/10.56294/mw20226>

**Enviado:** 02-07-2022

**Revisado:** 25-09-2022

**Aceptado:** 23-11-2022

**Publicado:** 24-11-2022

**Editor:** Dr. José Alejandro Rodríguez-Pérez 

### ABSTRACT

In the healthcare sector, Information Systems are fundamental for decision-making at all levels. This process can be strengthened with the implementation of Big Data analytics. In this context, the present study aims to describe the experiences, benefits, and applications of Big Data in Health Information Systems through a systematic literature review. The research reviewed 22 studies on the use of Big Data in the healthcare sector, applying inclusion and exclusion criteria to select relevant studies. The results of these studies indicated that the use of Big Data in healthcare can improve the quality of teleassistance services for patients, as well as logistics and financial services. It can also prevent diseases and improve patient care in public health information systems. Additionally, it can provide new knowledge and actionable information from new data sources, and promote the natural transformation of descriptive research into predictive and prescriptive research. The studies also highlighted the importance of Artificial Intelligence for data to be useful for research and medical purposes. In general, the research concluded that Big Data has a favorable impact on the healthcare sector, especially in biomedical research, and that its use can improve the efficiency and quality of medical care.

**Keywords:** Big Data; Data Analysis; Health; Health Information Systems.

### RESUMEN

En el sector de la salud, los Sistemas de Información son fundamentales para la toma de decisiones en todos los niveles. Este proceso se puede fortalecer con la implementación de los análisis de Big Data. En este contexto, el presente estudio se propone describir las experiencias, beneficios y aplicaciones del Big Data en los Sistemas de Información en Salud, a partir de una revisión sistemática de la literatura. La investigación revisó 22 estudios sobre el uso de Big Data en el sector de la salud. Se aplicaron criterios de inclusión y exclusión para seleccionar los estudios relevantes. Los resultados de dichos estudios indicaron que el uso de Big Data en la salud puede mejorar la calidad de los servicios de teleasistencia en pacientes, así como en los servicios logísticos y financieros. También puede prevenir enfermedades y mejorar la atención de los pacientes en los sistemas de información en salud pública. Además, puede proporcionar nuevos conocimientos e información procesable a partir de nuevas fuentes de datos, y propiciar la transformación natural de la investigación descriptiva en predictiva y prescriptiva. Los estudios también destacaron la importancia de la Inteligencia Artificial para que los datos puedan ser útiles para la investigación y los usos médicos. En general, la investigación concluyó que el Big Data tiene un impacto favorable en el sector de la salud, especialmente en la investigación biomédica, y que su uso puede mejorar la eficiencia y la calidad de la atención médica.

**Palabras clave:** Análisis de datos; Big Data; Salud; Sistemas de Información en Salud.

## INTRODUCCIÓN

La información en los sistemas sanitarios es el sustento en la toma de decisiones en todos los niveles. Esto requiere de integridad e interrelación de la misma en forma de sistema, que permita capturar, depurar, almacenar, recuperar, actualizar y tratar este recurso, con la finalidad de aprovecharlo en función de optimizar el funcionamiento de la atención sanitaria.<sup>(1)</sup>

Los Sistemas de Salud, en el cumplimiento de la misión por la que son concebidos generan y requieren de un constante flujo informativo. Este permite mantener un alto nivel de conocimiento de las actividades que se realizan en todos los niveles de atención, para la conducción de procesos directivos en los servicios. De esta manera, se ve reflejado el flujo de información, aportado por los Sistemas de Información en Salud (SIS), que requieren los decisores para la gerencia.<sup>(1,2)</sup>

A través de acciones y operaciones realizadas desde conocimientos previos en el área a administrar, los SIS propician el proceso íntegro de gestión de información para facilitar a los decisores las herramientas para abordar los problemas de salud de las poblaciones.<sup>(2)</sup> No obstante, lo ideal es que esta información cumpla con los criterios de veracidad, integridad y oportunidad para el aprovechamiento óptimo de la misma en las decisiones informadas.

Ante el reto que supone lidiar con un elevado volumen de datos en salud, las herramientas informáticas aportan competitividad al proceso de gestión de los SIS, con un reconocido y generalizado impacto positivo.<sup>(3)</sup> Las orientaciones de la Organización Mundial de la Salud (OMS) en este sentido abarcan un concepto amplio que implica el uso de las TIC en los sistemas de vigilancia, prevención, promoción y atención a la salud, la educación, los conocimientos y las investigaciones; y fomenta el desarrollo de planes estratégicos para la implantación de infraestructuras tecnológica en los servicios de salud.<sup>(4,5)</sup>

Este desarrollo conlleva a un nuevo enfoque y estudio de la realidad en los procesos de toma de decisiones. La implementación de tecnologías disruptivas en los SIS, fortalecen el proceso de transformación digital de la sociedad. Esto propicia cambios culturales en el campo de las investigaciones y estudios globales y particulares. Por tanto, es necesario adoptar experiencias, métodos y herramientas para la implementación y el desarrollo.<sup>(6)</sup>

Un ejemplo del impacto de las tecnologías en función de los SIS tuvo lugar durante la pandemia de COVID-19, donde fue necesaria de manera constante el flujo de información para una rápida toma de decisiones en los diferentes escenarios sanitarios y de la vida en general. Esta experiencia ha supuesto un cambio de paradigma en la aplicación de procedimientos tecnológicos en los Sistemas de Información en Salud, dada la cantidad de proyectos de transformación digital que se llevan a cabo.<sup>(6)</sup>

Estas transformaciones en los modos de gestionar información, y por consiguiente de administrar los sistemas sanitarios, se han vinculado estrechamente con la ciencia de datos. Algunos autores lo califican como dos elementos indisolubles, y se considera que la informatización ya no es posible sin la aplicación de la ciencia de datos. Se requiere la imbricación de tecnologías digitales y cultura organizacional, para mejorar o reemplazar procesos, con la finalidad de lograr un ecosistema de trabajo favorable que tribute a la calidad de la atención sanitaria.<sup>(6)</sup>

La ciencia de datos consiste en el estudio de datos, a fin de extraer e interpretar información significativa. Es un campo interdisciplinario emergente basado en la estadística, la informática, la computación, la comunicación, la gestión y la sociología para estudiar los datos y su entorno. Además, incluye la inteligencia artificial (IA), la minería de datos y las técnicas de *machine learning* o aprendizaje automático, entre otras; las cuales contribuyen al análisis de grandes cantidades de datos. Un análisis en profundidad de dichos datos los transforma en conocimientos, decisiones y acciones, a través de metodologías determinadas.<sup>(6,7,8,9)</sup>

El manejo de información desde estas perspectivas, contribuye al incremento exponencial de datos, y por tanto la necesidad de procesamiento e integración de una manera superior. Esto introduce el concepto de “datos masivos” o *Big Data*, el cual ha incorporado nuevas formas de análisis de la realidad a través de la aplicación de tecnologías disruptivas basadas en la ciencia de datos, que fortalecen el proceso de transformación digital.<sup>(6)</sup>

*Big Data* es el conjunto de herramientas para el trabajo con volúmenes de datos muy elevados de diferentes tipos y provenientes de diversas fuentes, a una velocidad gran velocidad. Con el apoyo de tecnologías emergentes es posible recopilar, almacenar, vincular y procesar dichos conjuntos de datos.<sup>(10)</sup>

Se han considerado muy positivas las aplicaciones del *Big Data* a través de modelos de *deep learning* en la mejoría de la salud en los ciudadanos, desde la información y la eficiencia de los sistemas sanitarios.<sup>(10,11)</sup> Ofrece prestaciones en la medicina preventiva, poblacional, personalizada, participativa, predictiva, la vigilancia de enfermedades infecciosas y las campañas de salud ambiental y laboral.<sup>(12)</sup>

Sin embargo, de forma similar a otras tecnologías disruptivas (*Blockchain*, tecnología NFT, IA, etc.) el *Big Data* está sujeto a debates acerca de la ética en la implementación del mismo.<sup>(13,14)</sup> A pesar de resaltarse el valor de *Big Data* en la sanidad, un grupo de científicos se cuestionan de qué manera se puede capturar el valor de las bases de datos masivos utilizados en esta tecnología de forma que se respeten los principios fundamentales de la ética y la privacidad.<sup>(10,15)</sup>

Se argumenta que las investigaciones con *Big Data* presentan nuevos riesgos, que no son abordables por las

normas regulatorias existentes. Es recomendable adaptar estas normas para fundamentar el uso de Big Data en principios éticos, normas comunitarias en investigación y expectativas del sujeto humano.<sup>(10)</sup>

Aunque existen bases de conocimiento en torno a los avances de la aplicación del Big Data en la salud y biomedicina, según se expande esta tecnología se identifican nuevos retos a los que enfrentarse, así como nuevas oportunidades que acrecientan el interés por el desarrollo de la investigación en este dominio.<sup>(16)</sup>

Dado los relieves que adquiere esta tecnología emergente por sus beneficios y los retos que supone la implementación en la salud humana, resulta importante identificar las tendencias y los desafíos del Big Data en los SIS. En este sentido, una revisión sistemática de literatura representa una contribución al conocimiento sobre esta temática, ante la necesidad de la evaluación crítica de la misma para el desarrollo de políticas de uso en el ámbito al que se hace referencia.

En correspondencia con lo anterior, el presente estudio se propone el objetivo de describir las experiencias, beneficios y aplicaciones del Big Data en los SIS, en el análisis de datos clínicos, epidemiológicos, o de gestión administrativa.

## MÉTODOS

Se realizó una revisión sistemática de literatura mediante la aplicación de la metodología establecida por PRISMA<sup>(17,18,19,20)</sup> (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) para este tipo de estudio. Con la revisión se pretende sintetizar información científica publicada acerca de los resultados de la implementación del Big Data en los SIS, desde el análisis de los beneficios de la implementación y los casos de éxitos reseñados en la literatura.

Para dicho propósito se realizó una búsqueda de información en diferentes sistemas de bases de datos bibliográficas, sistemas de indexación y resumen, y motores de búsqueda. (tabla 1). Los resultados obtenidos en cada uno de estos sistemas de bases de datos fueron refinados mediante filtros para simplificar la cantidad de resultados a una cifra considerable para una revisión más detallada de las fuentes recuperadas.

**Tabla 1.** Bases de datos consultadas y estrategias de búsqueda empleadas

Bases de datos	Estrategia de búsqueda
Scopus	( ALL ( "health information systems" ) AND ALL ( "big data" ) ) OR ( SUBJAREA ( medi OR nurs OR vete OR dent OR heal OR mult ) AND ALL ( "big data" ) ) AND ( LIMIT-TO ( SRCTYPE , "j" ) ) AND ( LIMIT-TO ( OA , "all" ) )
SciELO	(Big Data) AND (Sistemas de Información en Salud) /
Redalyc	(Big Data) AND (Health Information Systems)
PubMed/Medline	((Big Data) AND (Sistemas de Información en Salud)) OR ((Big Data) ) AND (Health Information Systems))
Google Scholar	

Al refinar los resultados de búsqueda en las bases de datos antes citadas, y luego de descartar artículos duplicados, se consideraron para la revisión 160 estudios, de los cuales se comprobó el cumplimiento de los criterios de inclusión y exclusión expuestos a continuación:

### Criterios de inclusión:

- Artículos sobre las aplicaciones, beneficios, experiencias y casos de éxito de la implementación del Big Data en la gestión de datos de los Sistemas de Información en Salud.
- Investigaciones que tratan el Big Data en el análisis de datos clínicos, epidemiológicos, o de gestión administrativa; independiente del tipo de diseño o enfoque del estudio.
- Documentos disponibles en línea de acceso abierto.
- Idioma: inglés o español.
- Publicados entre 2016 y 2022.

### Criterios de exclusión:

- Investigaciones asociadas a otros sectores ajenos a la salud y sin puntos de contacto con la misma dentro del estudio.
- Artículos abordados desde una perspectiva demasiado personal y poco práctica.
- Artículos que sean de pago.
- Artículos que no cumplen con el periodo de búsqueda.

Después de aplicar los criterios de inclusión y exclusión, se ha realizado una restricción en la muestra con el objetivo de analizar únicamente aquellos artículos que proporcionen información relevante para el objetivo

planteado.

Según dichos criterios y con la lectura del título fueron descartados 82 artículos en la etapa de cribado por no guardar relación directa con los objetivos del presente estudio, para un total de 78 que permanecieron en el proceso de evaluación. Con la lectura del resumen y del texto completo se descartaron 60 por salirse del contexto de la salud o carecer de resultados relevantes para reseñar en el estudio. En consecuencia, se incluyeron 22 estudios en la revisión (figura 1).

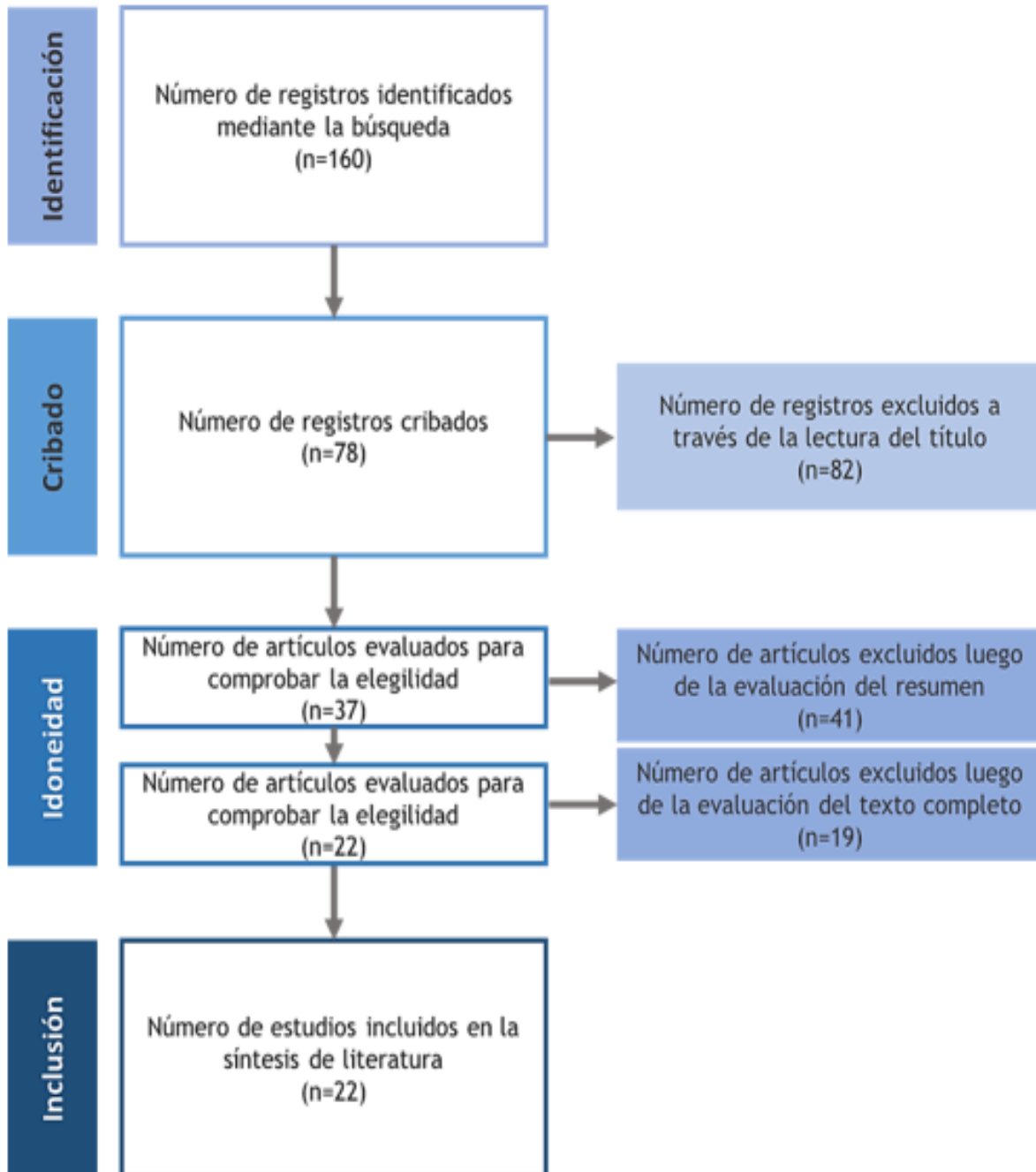


Figura 1. Diagrama de flujo de la búsqueda y selección de estudios según la metodología PRISMA

## RESULTADOS Y DISCUSIÓN

Se realizó la distribución según año de publicación de los estudios incluidos en la presente revisión. Del total de estudios analizados, fueron más frecuentes los publicados en el año 2020 con nueve documentos (40,9 %). A este le secundó el año 2022, con el 22,7 % de las publicaciones incluidas. El año 2018 no tuvo frecuencia en esta revisión (tabla 2).

Año de publicación	No.	%
2016	1	4,55
2017	2	9,09
2018	0	0,00
2019	3	13,64
2020	9	40,91
2021	2	9,09
2022	5	22,73
Total	22	100,00

La distribución geográfica de los documentos seleccionados en relación con su país de origen se expone en los hallazgos obtenidos a través de esta revisión. Fueron más frecuentes los estudios provenientes de América del Sur (50 %), y de esta región se incluyeron tres estudios de Cuba (13,6 %); dos de Venezuela, Ecuador y Colombia, para un 9,1 % cada país; y Perú y México tuvieron un estudio cada uno (4,5 %). Los países de Europa tuvieron una representación del 31,8 %; Asia del 13,6 %, representado por China; y América del Norte, representado por Estados Unidos solo tuvo un estudio (4,5 %) (figura 2).

Estos hallazgos dejan en manifiesto la ausencia de homogeneidad en la distribución de las publicaciones para la temática en el área de la salud, en las diversas regiones o continentes analizados, a pesar de que las estrategias de búsqueda se diseñaron de forma que pudiera abarcar cualquier área geográfica.



Figura 2. Distribución según lugar de publicación de los estudios incluidos en la revisión sistemática

La tabla 3 presenta las características de los estudios en el contexto de la revisión sistemática. Para ello se han considerado los siguientes atributos: autoría, ubicación geográfica del lugar de publicación, metodología, y principales resultados y conclusiones.

En relación a la metodología empleada para los estudios, predominó el enfoque cualitativo, para un 63,6 %. El enfoque cuantitativo y mixto tuvieron iguales proporciones (18,2 %). Cabe destacar, que se esperaban mayores contribuciones con enfoque cuantitativo o mixto de los países más desarrollados. Sin embargo, los países europeos, así como Estados Unidos no aportaron estudios cuantitativos a la presente revisión.

Predominaron los estudios descriptivos y revisiones de literatura (40,9 % cada uno). En una menor proporción se analizaron estudios experimentales (9,1 %), analíticos (4,5 %) y de implementación (4,5 %). La mayor parte de los estudios descriptivos se obtuvieron de los países de la región europea.

Tabla 3. Características de los estudios analizados en la revisión sistemática

Autor	Año	País/ Región	Metodología	Resultados y conclusiones
Araque González et al. <sup>(21)</sup>	2021	Venezuela	Estudio descriptivo, enfoque cualitativo	Los resultados evidencian el impacto y la relevancia del Big Data y sus aplicaciones en diversos sectores industriales, incluyendo el de la sanidad. Se concluye que la integración de hardware y software en el campo del Big Data es indispensable para mejorar la calidad de los servicios de teleasistencia en pacientes, así como en los servicios logísticos y financieros.
Aceto et al. <sup>(22)</sup>	2020	Italia	Enfoque sistemático en la exploración de la literatura	Este trabajo es una referencia para investigadores y profesionales en TIC que buscan comprender la relación entre Big Data y la Industria 4.0, y cómo se aplican en la Historia Clínica 4.0 (HC 4.0). Big Data aporta nuevos conocimientos e información procesable a partir de nuevas fuentes de datos, y propicia la transformación natural de la investigación descriptiva en predictiva y prescriptiva. Este estudio recomienda aplicar las experiencias de éxito para abordar las necesidades del sector de la salud y de los sistemas de información y automatización de la salud.
López-Capote et al. <sup>(23)</sup>	2017	Colombia	Revisión de literatura	El uso de herramientas Big Data permite pronosticar y prevenir infecciones o enfermedades mediante componentes analíticos. Con millones de datos capturados desde diversas fuentes, almacenados en repositorios y procesados con diferentes algoritmos, se logra mejorar la atención de los pacientes en los sistemas de información en salud pública, con aumento de los niveles de eficiencia en todos los procesos de atención. Estas herramientas son útiles no solo a nivel de salud, sino también a nivel gubernamental. Sin embargo, su uso también requiere competencias investigativas asociadas a las políticas de privacidad, seguridad y leyes de la información.
Lemus-Delgado et al. <sup>(24)</sup>	2020	México	Enfoque cualitativo, análisis crítico de información	Las herramientas tecnológicas basadas en Big Data pueden enriquecer la comprensión sobre los fenómenos globales, siempre y cuando se asuma una actitud crítica que reconozca que tanto la elección de los datos como su análisis están embebidos en contextos históricos y sociales. Por lo tanto, como cualquier otro enfoque, la ciencia de datos representa una manera parcial de capturar y analizar información.
Colombo et al. <sup>(25)</sup>	2020	Europa	Estudio descriptivo, enfoque mixto	<p>El sector de la salud está atrasado en el uso de datos de alta calidad para apoyar una transformación digital en la atención médica. Existen tecnologías para armonizar, compartir, acceder y utilizar datos de salud, incluso los muy grandes y diversos volúmenes de Big Data que produce el sector. Para una transformación digital positiva, se necesitan sistemas de información armonizados, habilidades y actitudes adecuadas, marcos éticos y partes interesadas comprometidas.</p> <p>Los países líderes de la Organización para la Cooperación y el Desarrollo Económico (OCDE) se están beneficiando de los datos y las tecnologías digitales para mejorar los resultados y las experiencias de los pacientes, la calidad y el desempeño de la atención médica, la investigación y la innovación.</p> <p>La transformación digital en la atención médica, requiere una estrategia de salud digital que incluya la gobernanza de datos de salud; sistemas nacionales de registros médicos electrónicos estandarizados y coherentes; e incentivos y asignaciones que respalden la capacidad del personal sanitario y los pacientes para beneficiarse de la transformación digital.</p>

7 Araujo Inastrilla CR

Dimitrov <sup>(26)</sup>	2016	Bulgaria	Estudio descriptivo, enfoque cualitativo	<p>El Big Data en la salud es fundamental para la transformación digital de la atención médica. Los dispositivos portátiles y las aplicaciones móviles respaldan el ejercicio físico, la educación sanitaria, el seguimiento de los síntomas y la gestión colaborativa de enfermedades y la coordinación de la atención. Todos esos análisis de plataforma pueden aumentar la relevancia de las interpretaciones de los datos, reducir la cantidad de tiempo que los usuarios finales dedican a reconstruir los resultados de los datos.</p> <p>Los conocimientos obtenidos del análisis de Big Data impulsan la disrupción digital del mundo de la salud, los procesos comerciales y la toma de decisiones en tiempo real. Surge una nueva categoría de asesores de Salud Digital con habilidades y la capacidad para interpretar y comprender datos de salud y bienestar.</p>
Mas-Bermejo et al. <sup>(27)</sup>	2022	Cuba	Investigación de implementación con un enfoque mixto transdisciplinar.	<p>Se evaluó el impacto de las acciones de control del COVID-19, mediante el análisis clínico-epidemiológico, el uso de técnicas de Big Data de la movilidad y la estimación en tiempo real del número reproductivo efectivo. Los modelos multinivel y de inteligencia artificial permitieron comparar el comportamiento de la pandemia en Cuba y el mundo.</p> <p>Además, se propusieron indicadores para el des-escalado de las acciones y se evaluó la equidad en la respuesta del país ante el COVID-19. La investigación epidemiológica realizada con el apoyo de la modelación y los sistemas de información fue fundamental para la toma de decisiones del Gobierno y el Ministerio de Salud Pública en el enfrentamiento efectivo a la pandemia.</p>
Cotino-Hueso <sup>(28)</sup>	2020	España	Enfoque cualitativo, análisis crítico de información	<p>El Big Data y la IA tienen un impacto favorable en el sector de la salud, especialmente en la investigación biomédica. En el área de la sanidad, cada vez es más importante el procesamiento de datos personales relacionados con la salud en los sectores público y privado mediante herramientas digitales.</p> <p>Los datos que se tratan son estructurados, semiestructurados y, mayoritariamente, no estructurados y brutos, procedentes de diversas fuentes como sensores, registros médicos electrónicos, datos biométricos, historias clínicas, imágenes, pruebas, publicaciones, webs y redes sociales. La IA es esencial para que estos datos puedan ser útiles para la investigación y los usos médicos. En las acciones frente a la COVID-19 se ha de facilitar el flujo de estos datos entre los sectores público y privado.</p>
Prado-Ortega et al. <sup>(29)</sup>	2020	Ecuador	Estudio descriptivo, enfoque cuantitativo	<p>El Big Data tuvo un impacto significativo en la lucha contra la pandemia del COVID-19. Los instrumentos tecnológicos digitales utilizados por los sistemas sanitarios, como los mecanismos de medición termográfica y los sistemas de posicionamiento de georreferencia, permitieron analizar y segmentar la información de aglomeración de personas a través de mapas de calor geográfico. Esto permitió a las autoridades sanitarias utilizar esta información y difundirla en medios digitales para tomar medidas correctivas necesarias y dar respuestas en el menor tiempo posible. Además, se pudo localizar la aglomeración de personas no solo por sus datos corporales sino por el uso de datos. Las decisiones se apoyaron en estas herramientas tecnológicas y su análisis fue muy objetivo para brindar alternativas que permitieran minimizar los efectos de eventos como el dado por la pandemia del COVID-19.</p>

Chaochao et al. <sup>(30)</sup>	2020	China	Estudio analítico	El estudio analiza la aplicabilidad de varios algoritmos, proporcionando una referencia metodológica para que los laboratorios clínicos establezcan intervalos de referencia mediante la utilización de Big Data. A partir del estudio de tres conjuntos de datos, se encontró que la coherencia entre los diferentes algoritmos en los datos del examen físico era mayor que la de los datos de pacientes ambulatorios. El estudio propone el uso de algoritmos que se basan en datos de exámenes físicos para establecer intervalos de referencias para hormonas relacionadas con la tiroides en adultos mayores, y ofrece diferentes perspectivas según las características de los conjuntos de datos.
Viteri-Gonzales et al. <sup>(31)</sup>	2022	Perú	Investigación aplicada, pre-experimental con pre-prueba y post-prueba.	Se analizó la información de los datos abiertos del Ministerio de Salud y del Centro Nacional de Abastecimiento de Recursos Estratégicos en Salud, para el monitoreo y control de la emergencia sanitaria COVID-19 bajo el ecosistema de <i>Apache Hadoop</i> y <i>Microsoft Azure</i> . El análisis de la situación de emergencia llegó a determinar la necesidad de información resumida en 10 requerimientos para poder mantener un apoyo en la toma de decisiones sobre morbi-mortalidad por COVID-19.  El análisis de estos requerimientos llevó a consolidar un modelo estrella constelación, y por la cantidad de datos que se insertaron se propuso una arquitectura basada en la nube de <i>Microsoft Azure</i> . La arquitectura de Big Data en esta nube permitió una escalabilidad en servicios, pero en esta solución se configuraron los principales servicios para el soporte a los requerimientos. El <i>dashboard</i> creado en <i>Power BI</i> permitió visualizar de forma dinámica la información.
Changmarín <sup>(32)</sup>	2021	Venezuela	Enfoque cualitativo, consulta de expertos	La automatización en base a sistemas de Big Data propicia cambios en los puestos de trabajo, ya que los equipos inteligentes harán el trabajo rutinario de la contabilidad, a la vez que puede crear nuevos puestos para personal especializado en la gestión de estos sistemas. Los investigadores contarán con mayor información para proyectar sus estudios, en las áreas de modelado, análisis estadístico y minería de textos.  Las bases de datos relacionales distribuidas resultan más eficientes y más seguras contra el robo de datos u otros usos perjudiciales. Por lo tanto, se requieren procedimientos que permitan controlar a la captura de los datos, provenientes de actividades en el mundo físico.
Zerega-Prado et al. <sup>(33)</sup>	2022	Ecuador	Método empírico analítico, cuasi experimental con enfoque cuantitativo	Se desarrolla un modelo a través de una arquitectura en Big Data, de un sistema computacional que consiste en 4 capas: Capa Fuentes de Datos, Capa Transformación, Capa Herramientas y Capa Visualización, para una empresa ecuatoriana de seguros de salud. El modelo está orientado a la consolidación de la información.  Las distintas fuentes de información o tipos de datos utilizados en el ámbito de los seguros de salud permiten conocer varias aristas del análisis de datos, además de obtener indicadores para mejorar la toma de decisiones. Se concluye que el modelo es aplicable a empresas nacionales o extranjeras mediante la contrastación de factores viables en una empresa específica del medio.



Peng-Ting et al. <sup>(34)</sup>	2020	China	Revisión de literatura	<p>El sistema de información sanitaria basado en Big data ha experimentado un crecimiento significativo y se está adaptando a la información médica para derivar tendencias de salud importantes y respaldar la atención preventiva y oportuna. Sin embargo, las organizaciones enfrentan barreras para implementar un sistema de información sanitaria basado en Big Data.</p> <p>Para superar estas barreras, se puede adoptar el enfoque del proceso de red analítica para determinar el peso del aspecto y aplicar el método VIKOR para resolver problemas de decisión con criterios en conflicto, y de esta forma concluir una estrategia apropiada.</p> <p>El modelo propuesto puede proporcionar a los gestores hospitalarios previsiones e implicaciones que faciliten la eliminación de barreras organizativas a la hora de adoptar el SIS basado en Big Data. Al comprender la secuencia de importancia de los factores de resistencia, los gerentes pueden formular estrategias eficientes para resolver problemas con prioridades apropiadas.</p>
Gutiérrez-Martínez et al. <sup>(35)</sup>	2019	Cuba	Revisión de literatura	<p>A través de las técnicas analíticas, el Big Data en medicina permite conocer y evaluar los riesgos, prevenir los accidentes o decidir cuál es el mejor tratamiento para un paciente, además de mejorar la gestión en los hospitales. La recopilación de grandes bases de información ha permitido el avance significativo en la prevención y el diagnóstico.</p> <p>La interconexión permanente de profesionales sanitarios a nivel mundial ha facilitado que el diagnóstico se pueda realizar desde parámetros más amplios a los que se podía acceder antes que apareciera este tipo de tecnología. El Big Data se convierte en una oportunidad sin precedentes de accesibilidad a una masiva cantidad de datos desbloqueados. Esta oportunidad pretende que para 2030, exista información más contextualizada, diagnósticos más concretos y atención más personalizada.</p>
De Lecuona <sup>(36)</sup>	2020	España	Enfoque cualitativo, análisis de información	<p>En el contexto de la pandemia por COVID-19, se incrementó el uso de la inteligencia artificial, el Big Data y las aplicaciones móviles en el sector de la salud. Sin embargo, la tendencia a la mercantilización de datos personales y la "economía de la atención" promovida por las grandes empresas tecnológicas, así como la discriminación algorítmica y la acumulación indiscriminada de datos, plantean cuestiones importantes sobre la libertad y la intimidad. El análisis filosófico-jurídico de estas cuestiones requiere una comprensión previa de los aspectos técnicos complejos para construir una reflexión sólida sobre las cuestiones bioéticas de las tecnologías emergentes.</p>
Dash et al. <sup>(37)</sup>	2019	Portugal	Enfoque mixto, estudio descriptivo desde el análisis de información	<p>En la industria de la salud, varias fuentes de Big Data incluyen registros hospitalarios, registros médicos de pacientes, resultados de exámenes médicos y dispositivos que forman parte del Internet de las cosas. La investigación biomédica también genera una parte importante de Big Data relevantes para la atención sanitaria pública. Sin embargo, estos datos requieren una gestión y un análisis adecuados para obtener información significativa. Existen varios desafíos asociados con cada paso del manejo de Big Data que solo pueden superarse mediante el uso de soluciones informáticas de alta gama. Una gestión, análisis e interpretación eficientes de Big Data pueden abrir nuevas vías para la atención sanitaria moderna con mejores servicios y ventajas financieras.</p>

Zhihan et al. <sup>(38)</sup>	2020	China	Revisión de literatura	En el contexto del Big Data, se analizan los factores de riesgo de privacidad y seguridad en los datos médicos en China. Los resultados del cuestionario muestran que la probabilidad de análisis de datos, proceso de tratamiento médico, proceso de diagnóstico de enfermedades, falta de medidas de protección y sistema de acceso imperfecto es superior a 0,8 cuando los Big Data de atención médica están orientados a servicios en la nube. En base a esto, se proponen dos niveles de medidas de protección de la privacidad: tecnología y gestión. Las instituciones médicas deben prestar atención a la protección de la privacidad de los datos y aprovechar el uso de datos médicos digitales para brindar apoyo a las decisiones para el análisis posterior de datos médicos.
Menasalvas et al. <sup>(39)</sup>	2017	España	Enfoque cualitativo, análisis de información	El Big Data en la atención médica supondrá una mejora significativa en la calidad de la atención a los pacientes, así como en la reducción notable en los costes de sanidad. Para alcanzar estos logros es fundamental la integración de todos los datos procedentes de muy diferentes fuentes, así como el desarrollo de nuevas tecnologías que permitan la explotación de dichos datos. Sin embargo, el verdadero valor de Big Data en salud se conseguirá solo si los diferentes actores implicados en el proceso se comprometen en este proyecto de forma conjunta para llevar el ámbito sanitario a una nueva era.
Vanegas-Casadiegos <sup>(40)</sup>	2022	Colombia	Estudio descriptivo, enfoque mixto	Los costos son un componente importante en la mejora de los procesos de atención en salud, y están estrechamente relacionados con la racionalidad y eficacia en los recursos que se invierten en la atención. Los incidentes y eventos adversos presentes en cualquier fase de la prestación del servicio impactan los principios generales del sistema en salud. Por lo tanto, es importante integrar los resultados de la transformación de datos realizados a través de herramientas de Big Data con el análisis multidisciplinario y variables clínicas para determinar la frecuencia de asociación entre las variables y los factores de riesgo que impactan dichos resultados. Las conclusiones deben ser integradas a los planes de mejora de la institución, buscando las acciones pertinentes en la disminución de estos eventos adversos, lo que se traduce en un menor gasto y mejora en la calidad de la atención en salud.
Santos-Domínguez <sup>(41)</sup>	2022	Cuba	Mapeo sistemático	Es prioritario utilizar el Big Data en beneficio de los ensayos clínicos. En Cuba no hacen uso de los datos almacenados previamente para mejorar el diseño de los mismos. Se observa que el uso de técnicas de minería de datos y aprendizaje automático ayuda a optimizar el diseño de ensayos clínicos, lo que puede significar ahorros millonarios en la industria biofarmacéutica.
Nicholson-Price et al. <sup>(42)</sup>	2019	Estados Unidos	Revisión de literatura	El Big Data se ha convertido en la consigna omnipresente de la innovación médica. El rápido desarrollo de las técnicas de aprendizaje automático y de la inteligencia artificial en particular ha prometido revolucionar la práctica médica, desde la asignación de recursos hasta el diagnóstico de enfermedades complejas. Sin embargo, entraña riesgos y desafíos, entre ellos importantes cuestiones sobre la privacidad del paciente.

Los resultados de los estudios de manera general muestran el potencial del Big Data en la mejoría de la calidad de la atención sanitaria, al proveer a los SIS con herramientas y metodologías que permiten tener una visión más detallada de los servicios de salud. Entre los beneficios que reseña están la reducción de costos mediante la identificación de patrones de enfermedades que requieren de tratamientos encarecidos; además de las ventajas en la toma de decisiones informadas y en poblaciones muy amplias.

Con relación a la práctica clínica se destacan los valores que añade el Big Data a la misma. La integración de diferentes fuentes de información disponibles facilita el objetivo de diagnosticar y tratar a los pacientes con una base de conocimientos más sólida y personalizada.<sup>(21,22,23,24,25,26)</sup> La combinación con técnicas de IA permite aplicar una capa de analítica e inteligencia a la información que el personal médico maneja, a través de sistemas de historia clínica digital.<sup>(22,33)</sup> Se demuestra que Big Data tiene cabida en ámbitos como la genómica, los ensayos clínicos, la operativa clínica, la colaboración ciudadana o la teleasistencia.<sup>(28,30,35,41,42)</sup>

En el análisis de datos epidemiológicos, ha destacado en las investigaciones médicas a gran escala, la predicción de brotes de enfermedades, la identificación de factores de riesgo y en la planificación de medidas de prevención.<sup>(27,29,31,36,41)</sup> Esto ha sido posible con el apoyo de sistemas de geolocalización y georreferencia. Los principales esfuerzos en este sentido se enmarcan en el contexto de la pandemia de COVID-19, donde diferentes sistemas de salud tomaron ventaja del Big Data en el monitoreo de la morbilidad y mortalidad del virus, para actuar en consecuencia de forma rápida y acertada.<sup>(27,28,29,31)</sup>

Algunos estudios refieren la aplicabilidad en el análisis de datos para la gestión administrativa de la salud.<sup>(27,32,33,34,37)</sup> Los análisis de datos masivos ofrecen una perspectiva mucho más efectiva para la concepción de políticas, planes de mejora y estrategias a seguir por las instituciones sanitarias. Se optimiza la toma de decisiones con impacto económico y social significativo. Los modelos predictivos que anticipen las necesidades sanitarias coadyuvan a la gerencia desde la asignación eficiente de recursos.

Sin embargo, fueron evidentes las preocupaciones en torno a las problemáticas de tipo éticas y legales inherentes a los análisis de Big Data, por las insuficiencias del marco regulatorio acerca de los principios éticos en su uso, normas comunitarias en investigación y expectativas del sujeto humano. Esto muestra que el uso de Big Data en los SIS debe ser cuidadosamente regulado para proteger la privacidad y seguridad de los datos médicos.<sup>(33,39,40)</sup>

Además, se hace alusión a las barreras organizativas que enfrentan las instituciones para implementar SIS basados o apoyados en Big Data. Esto indica que la implementación puede ser un proceso complejo que requiere medidas de seguridad adecuadas, cambios estructurales y culturales de los organismos implicados.<sup>(21,31,40)</sup>

## CONCLUSIONES

Tras el esfuerzo por sintetizar e integrar los hallazgos acerca de la implementación de Big Data en los SIS, se puede concluir que esta tecnología disruptiva supone un cambio de paradigma muy útil y necesario en los sistemas de salud. El proceso de generalización demanda superar la brecha tecnológica, organizativa y legal existente que no permite una explotación con ética de las bondades que ofrece a la atención médica eficiente, segura y personalizada.

## REFERENCIAS BIBLIOGRÁFICAS

1. Lazo-Pérez MA. La epidemiología de las pandemias. *Revista cubana de Tecnología de la Salud*. 2021;12(2):147-154.
2. Gutiérrez-Vera D. Los Sistemas de Información en Salud, relevancia y realidades ante la pandemia del virus SARS-COV-2. *Rev. CMV*. 2023;1(1-3):e007.
3. Preciado-Rodríguez AJ, Valles-Coral MA, Lévano-Rodríguez D. Importancia del uso de sistemas de información en la automatización de historiales clínicos, una revisión sistemática. *Revista Cubana de Informática Médica*. 2021;13(1):e417.
4. Organización Mundial de la Salud. Proyecto de estrategia mundial sobre salud digital. 2020-2025. OMS; 2020. [https://cdn.who.int/media/docs/default-source/documents/200067-lb-full-draft-digital-health-strategy-with-annex-cf-6jan20-cf-rev-10-1-clean-sp\\_1c8b2b9c-4c25-4efb-8553-9f466028b583.pdf?sfvrsn=4b848c08\\_4](https://cdn.who.int/media/docs/default-source/documents/200067-lb-full-draft-digital-health-strategy-with-annex-cf-6jan20-cf-rev-10-1-clean-sp_1c8b2b9c-4c25-4efb-8553-9f466028b583.pdf?sfvrsn=4b848c08_4)
5. Delgado-Ramos A, Vidal-Ledo M, Rodríguez-Díaz A, Barthelemy-Aguilar K, Torres-Ávila D. Salud y transformación digital. *Educación Médica Superior*. 2022;36(2):e3442. <https://ems.sld.cu/index.php/ems/article/view/3442>
6. Vidal-Ledo M, Delgado-Ramos A, Gutiérrez-Vera D, Rodríguez-Díaz A. Ciencia de Datos en Salud. *Educación Médica Superior*. 2023; 37(1):e3735. <https://ems.sld.cu/index.php/ems/article/view/3735>

7. Rosa JM, Frutos EL. Ciencia de datos en salud: desafíos y oportunidades en América Latina. *Rev. Med. Clin. Condes*. 2022;33(6):591-597. <https://doi.org/10.1016/j.rmclc.2022.09.007>
8. Ramos-Delgado A, Vidal-Ledo M, Rodríguez-Díaz A, Barthelemy-Aguilar K, Torres-Ávila D. Salud y transformación digital. *Educ Méd Super*. 2022;36(2). <http://www.ems.sld.cu/index.php/ems/article/view/3442>
9. Lemus-Delgado D, Pérez-Navarro R. Ciencia de datos y estudios globales: aportaciones y desafíos metodológicos. *Colomb. Int.* 2020;(102):41-62. [http://www.scielo.org.co/scielo.php?pid=S0121-56122020000200041&script=sci\\_abstract&tlng=es](http://www.scielo.org.co/scielo.php?pid=S0121-56122020000200041&script=sci_abstract&tlng=es)
10. Rumiche-Chávarry RP, Matas-Terrón A, Ríos-Ariza JM, Chunga-Chinguel R. Implicaciones sociales de la irrupción del Big Data y la robótica: un análisis prospectivo en docentes hispanoamericanos. *Revista Electrónica Interuniversitaria de Formación del Profesorado*. 2023;26(1):115-127. <https://doi.org/10.6018/reifop.543871>
11. Alcalde G, Alfonso I. Utilización de tecnología Big Data en investigación clínica. *Revista de Derecho y Genoma Humano*. 2019;(1 N° extraordinario):55-83. <https://dx.doi.org/10.14679/1133>
12. Sousa MJ, Pesqueira AM, Lemos C, Sousa M, Rocha A. Decision-Making based on Big Data Analytics for People Management in Healthcare Organizations. *Journal of Medical Systems*. 2019;43:290. <https://doi.org/10.1007/s10916-019-1419-x>
13. Araujo-Inastrilla CR, Vitón-Castillo AA. Blockchain in health sciences: Research trends in Scopus. *Iberoamerican Journal of Science Measurement and Communication*. 2023; 3(2), 1-10. <https://doi.org/10.47909/ijsmc.56>
14. Araujo-Inastrilla CR. La tecnología NFT y la información en salud: un debate ético contemporáneo. *Revista Cubana de Informática Médica*. 2023;15(1):e618. <https://revinformatica.sld.cu/index.php/rcim/article/view/618>
15. Blaconá MT. Reflexiones sobre los nuevos desafíos éticos que plantea el uso de grandes bases de datos (Big-Data) en investigación. *SaberEs*. 2019;11(2).
16. Parra-Calderón CL. Big data en sanidad en España: la oportunidad de una estrategia nacional. *Gac Sanit*. 2016;30(1):63-65. <http://dx.doi.org/10.1016/j.gaceta.2015.10.005>
17. Pagea MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Rev Esp Cardiol*. 2021;74(9):790-799 <https://doi.org/10.1016/j.recesp.2021.06.0>
18. Sánchez MJ. Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta*. 2010;38(2):53-64. <http://hdl.handle.net/11162/5126>
19. Pigott TD, Polanin JR. Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*. 2020;90(1):24-46. <https://doi.org/10.3102/0034654319877153>
20. Serrano SS, Navarro IP, González MD. ¿Cómo hacer una revisión sistemática siguiendo el protocolo PRISMA?: Usos y estrategias fundamentales para su aplicación en el ámbito educativo a través de un caso práctico. *Bordón: Revista de pedagogía*. 2022;74(3):51-66.
21. Araque-González GA, Gómez-Vásquez M, Vélez-Uribe JP, Suárez-Hernández AH. Big Data y las implicaciones en la cuarta revolución industrial - Retos, oportunidades y tendencias futuras. *Revista Venezolana de Gerencia*. 2021;26(93). <https://www.redalyc.org/articulo.oa?id=29066223003>
22. Aceto G, Persico V, Pescapé A. Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0. *Journal of Industrial Information Integration*. 2020;18. <https://doi.org/10.1016/j.jii.2020.100129>
23. Baker K, Dunwoodie E, Jones EG, Newsham A, Johnson O, Price CP, et al. Process Mining Routinely

Collected Electronic Health Records to Define Real-Life Clinical Pathways during Chemotherapy. *International Journal of Medical Informatics*. 2017. <http://dx.doi.org/10.1016/j.ijmedinf.2017.03.011>

24. Lemus-Delgado D, Pérez-Navarro R. Ciencias de datos y estudios globales: aportaciones y desafíos metodológicos. *Colombia Internacional*. 2022;102:41-62. <https://doi.org/10.7440/colombiaint102.2020.03>

25. Colombo F, Oderkirk J, Slawomirski L. Sistemas de Información en Salud, Electrónicos Salud: avances y desafíos en Registros médicos y Big Data en el mundo Países de la OCDE. *Manual de Salud Global*. Francia: OMS; 2020. [https://doi.org/10.1007/9783030053253\\_711](https://doi.org/10.1007/9783030053253_711)

26. Dimitrov DV. Medical Internet of Things and Big Data in Healthcare. *Healthc Inform Res*. 2016 July;22(3):156-163. <http://dx.doi.org/10.4258/hir.2016.22.3.156>

27. Mas-Bermejo P, Sánchez-Valdés L, Vidal-Ledo M, Seuc Jo A, et al. Contribuciones de la epidemiología, la modelación y los sistemas de información en el enfrentamiento de la COVID-19. *An Acad Cienc Cuba*. 2022;12(3):e1221. <http://www.revistaccuba.cu/index.php/revacc/article/view/1221>

28. Cotino-Hueso L. Inteligencia artificial, big data y aplicaciones contra la COVID-19: privacidad y protección de datos. *Revista de los Estudios de Derecho y Ciencia Política*. 2020;31:1-17. <http://dx.doi.org/10.7238/idp.v0i31.3244>

29. Prado-Ortega M, Grunauer-Robalino R. Salud pública: aplicación de Big Data en detección de concentración poblacional, para evitar brotes epidemiológicos por COVID-19. *Identidad bolivariana*. 2020;4(2).

30. Chaochao M, Yutong Z, Li'an H, Yicong Y, Fang Z, et al. Validación y comparación de cinco algoritmos de minería de datos utilizando big data de laboratorios clínicos para establecer intervalos de referencia de hormonas tiroideas para adultos mayores. *Clinical Biochemistry*. 2022;107:40-49. <https://doi.org/10.1016/j.clinbiochem.2022.05.008>

31. Viteri-Gonzales AP, Beltrán-García JA. Solución de Big Data para el análisis de los datos abiertos de MINSA y CENARES para el monitoreo y control de la emergencia sanitaria COVID-19 bajo el ecosistema de Apache Hadoop y Microsoft Azure. Trujillo: Universidad Privada de Antenor Orrego; 2022.

32. Changmarín CA. Big data y su impacto en el ejercicio de la contaduría pública, las empresas y los sistemas de información: Una mirada a la ética. *Actualidad Contable Faces*. 2021 Enero-Junio;24(42):9-35. <https://doi.org/10.53766/ACCON/2021.42.01>

33. Zerega-Prado J, Llerena-Izquierdo J. Arquitectura de consolidación de la información para seguros de la salud mediante Big Data. *Memoria Investigaciones en Ingeniería*. 2022;23:18-31. <https://doi.org/10.36561/ING.23.3>

34. Peng-Ting C, Chia-Li L, Wan-Ning W. Big data management in healthcare: Adoption challenges and implications. *International Journal of Information Management*. 2020. <https://doi.org/10.1016/j.ijinfomgt.2020.102078>

35. Gutiérrez-Martínez JA, Febles-Estrada A. Las tecnologías disruptivas y su aplicación en la medicina con vistas al 2030. *Revista Cubana de Salud Pública*. 2019;45(4):e1563.

36. De Lecuona I. Aspectos éticos, legales y sociales del uso de la inteligencia artificial y el Big Data en salud en un contexto de pandemia. *Revista Internacional de Pensamiento Político*. 2020;15:139-166.

37. Dash S, Shakyawar S, Sharma M, Kaushik S. Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*. 2019;6:54 <https://doi.org/10.1186/s40537-019-0217-0>

38. Zhihan L, Liang Q. Analysis of healthcare big data. *Future Generation Computer Systems*. 2020; 109:103-110. <https://doi.org/10.1016/j.future.2020.03.039>

39. Menasalvas E, Gonzalo C, Rodríguez-González A. Big data en salud: retos y oportunidades. *Economía industrial*. 2017;405:87-97. <https://dialnet.unirioja.es/servlet/articulo?codigo=6207516>

40. Vanegas-Casadiago JA. Big Data a path to Healthcare Quality. Bogotá: Universidad El Bosque; 2022.

41. Santos-Domínguez A. Big data y aprendizaje automático para mejorar los procesos en los Ensayos Clínicos: mapeo sistemático de la literatura. CubaSalud. 2022. <https://convencionalud.sld.cu/index.php/convencionalud22/2022/paper/download/2910/1195>

42. Nicholson-Price W, Glenn-Cohen I. Privacy in the age of medical big data. Nature Medicine. 2019;25:37-43. <https://doi.org/10.1038/s41591-018-0272-7>

#### **FINANCIACIÓN**

Los autores no recibieron financiación para el desarrollo de la presente investigación.

#### **CONFLICTO DE INTERESES**

Se declara que no existe conflicto de intereses.

#### **CONTRIBUCIÓN DE AUTORÍA**

*Conceptualización:* Carlos Rafael Araujo Inastrilla.

*Curación de datos:* Carlos Rafael Araujo Inastrilla.

*Análisis formal:* Carlos Rafael Araujo Inastrilla.

*Adquisición de fondos:* Carlos Rafael Araujo Inastrilla.

*Investigación:* Carlos Rafael Araujo Inastrilla.

*Metodología:* Carlos Rafael Araujo Inastrilla.

*Administración del proyecto:* Carlos Rafael Araujo Inastrilla.

*Recursos:* Carlos Rafael Araujo Inastrilla.

*Software:* Carlos Rafael Araujo Inastrilla.

*Supervisión:* Carlos Rafael Araujo Inastrilla.

*Validación:* Carlos Rafael Araujo Inastrilla.

*Visualización:* Carlos Rafael Araujo Inastrilla.

*Redacción - borrador original:* Carlos Rafael Araujo Inastrilla.

*Redacción - revisión y edición:* Carlos Rafael Araujo Inastrilla.