Rev Clin Esp. 2023;223(Supl 1):S143

Revista Clínica Española

https://www.revclinesp.es

1348 - ¿CONOCE EL *BIG DATA*? UNA ENCUESTA A MÉDICOS DE MEDICINA INTERNA

Jessica Wong Seoane[1] *y María Brey Devesa*[2]

[1]*Hospital Universitario de Araba, Vitoria-Gasteiz.* [2]*Hospital Clínic, Barcelona.*

Resumen

Objetivos: El *Big Data* es un término que ha venido para quedarse. Sin embargo, existe gran vaguedad conceptual, con una falta de definición formal aceptada que delimite su concepto, utilidad y funcionamiento[1]. El objetivo de este estudio fue evaluar el grado de conocimiento que poseen los médicos de Medicina Interna en España sobre el *Big Data*.

Métodos: Realizamos una encuesta mediante Google Forms entre los diferentes médicos de Medicina Interna de varios centros hospitalarios entre el 27 de junio y el 3 de julio de 2022. Las respuestas eran registradas de manera anónima y con posibilidad de una única respuesta. Como criterio de inclusión en el estudio, era obligatorio el consentimiento expreso por parte del participante. Excluimos aquellas respuestas de centros hospitalarios no españoles. Además, efectuamos un análisis estadístico descriptivo y un análisis bivariable de las respuestas registradas.

Resultados: Contamos con una muestra de 53 participantes de 20 centros hospitalarios de España. La muestra de encuestados constaba de 28 hombres (52,8%) y de 25 mujeres (47,2%). El rango de edad era de 24 a 67 años con una media de edad de 39 años. 14 participantes (26,4%) admitían desconocer qué era el *Big Data*, mientras que 39 (73,6%) afirmaron conocerlo. Únicamente 8 participantes (15,1%) afirmaban conocer bien su concepto y su funcionamiento. Entre los participantes había más hombres que mujeres que confirmaban conocer el *Big Data* de manera estadísticamente significativa ($p = 0,015$). No hubo diferencias entre la edad ($p = 0,214$) ni la categoría profesional ($p = 0,700$). Entre los encuestados, la definición más aceptada de *Big Data* fue la definición de la Comisión Europea («grandes cantidades de datos producidos muy rápidamente por un elevado número de fuentes diversas. Los datos pueden ser creados por personas o generados por máquinas, como los sensores que recogen información sobre el clima, las imágenes por satélite, las fotografías y vídeos digitales, los registros de transacciones de compra, las señales de GPS, etc. Abarca muchos sectores, desde la sanidad hasta el transporte o la energía»)[2]. En la definición proporcionada por los encuestados, coincidían en su mayoría en que el *Big Data* es «una gran cantidad de datos», «con gran precisión y similitud con la realidad», «recopilados a partir de medios digitales», «que se puede analizar mediante inteligencia artificial», «con el objetivo de obtener patrones y tendencias para poder hacer predicciones». En cuanto a formación, 10 participantes (18,9%) habían recibido previamente formación en *Big Data*, aunque 36 participantes (67,9%) «sí» estarían interesados y 16 participantes (30,2%) «tal vez» estarían interesados en tener formación en *Big Data*.

Variables	[ALL] N = 53	N
Edad	39,4 (11,8)	53
Sexo:		53
Hombre	28 (52,8%)	
Mujer	25 (47,2%)	
Categoría profesional:		53
Médico adjunto	23 (43,4%)	
Médico especialista	9 (17,0%)	
MIR 1	4 (7,55%)	
MIR 2	1 (1,89%)	
MIR 4	6 (11,3%)	
MIR 5	10 (18,9%)	

Variables	[ALL] N = 53	N
¿Sabe lo que es *Big Data*?:		53
No	14 (26,4%)	
Sí	39 (73,6%)	
¿Cuánto conoce sobre el *Big Data*?:		53
Creo entender bien su concepto y su funcionamiento	8 (15,1%)	
He escuchado sobre ello, pero no tengo claro su concepto ni su funcionamiento	27 (50,9%)	
No sé lo que es	7 (13,2%)	
Tengo claro su concepto, pero desconozco su funcionamiento	11 (20,8%)	
¿Tiene conocimiento si usan *Big Data* en el hospital donde trabaja?:		53
Están en proceso de utilizarlo	5 (9,43%)	
Lo desconozco	28 (52,8%)	
No	16 (30,2%)	
Sí	4 (7,55%)	
¿Ha participado en algún proyecto de investigación relacionado o con el uso de *Big Data*?:		53
Estamos en proceso de utilizarlo	1 (1,89%)	
No	48 (90,6%)	
Sí	4 (7,55%)	
¿Le han dado alguna formación relacionada con *Big Data*?:		53

No	43 (81,1%)	
Sí	10 (18,9%)	
¿Le gustaría tener formación en *Big Data*?:		53
No	1 (1,89%)	
Sí	36 (67,9%)	
Tal vez	16 (30,2%)	

Variables	No N = 14	Sí N = 39	OR	p, ratio	p, total
Edad	33,0 [28,5;44,2]	41,0 [28,5;49,5]	1,04 [0,98;1,11]	0,159	0,214
Sexo					0,015
Hombre	3 (21,4%)	25 (64,1%)	Ref.	Ref.	
Mujer	11 (78,6%)	14 (35,9%)	0,16 [0,03;0,64]	0,008	
Categoría profesional					0,700
Médico adjunto	7 (50,0%)	16 (41,0%)	Ref.	Ref.	
Médico especialista	1 (7,14%)	8 (20,5%)			
MIR 1	2 (14,3%)	2 (5,13%)			
MIR 2	0 (0,00%)	1 (2,56%)			
MIR 4	2 (14,3%)	4 (10,3%)			
MIR 5	2 (14,3%)	8 (20,5%)		,	

¿Cómo definiría el concepto *Big Data*?
1. Como una herramienta útil para unificar grandes volúmenes de información.
2. Algo así como el análisis de paquetes masivos de información recogidos de diferentes formas.
3. La informatización del sistema sanitario y su depósito automático en bases de datos facilita la extracción y análisis de datos de los pacientes.
4. El conjunto de datos virtuales disponibles a nivel mundial.
5. Obtención y análisis de datos masivos para extraer conclusiones y aplicarlos en mejorar distintos procesos.
6. Una gran base de datos que usa estos datos relacionándolos entre ellos para conseguir información para mejorar los objetivos a los que van encaminados.
7. Es información extraíble a través de redes sociales/consultas a través de internet/red de gran cantidad de población.
8. Un volumen grande de datos que necesitan de técnicas especiales para trabajar con ellos.
9. El análisis de datos masivos producidos por la actividad de las personas en la red.
10. Utilización de gran cantidad de datos de múltiples fuentes para la investigación científica.

11. Analizar datos a gran escala para identificar errores/mejoras y proponer cambios.
12. Análisis de grandes cantidades de datos para detectar tendencias.
13. Uso de grandes cantidades de datos.
14. Procesamiento de múltiples datos para obtener patrones.
15. Grandes cantidades de información.
16. Análisis de datos.
17. Conjunto de datos que se generan en el día a día de las actividades económicas y sociales.
18. Análisis de grandes cantidades de datos.
19. Conjunto acoplado de datos.
20. Es el conjunto global de datos de los que disponemos de un modo u otro (bases de datos, medios de comunicación, redes sociales...) qué se pueden utilizar para realizar estudios o estimaciones sobre un tema, con una gran precisión y similitud con la realidad, al tratarse de una muestra muy grande y de escala mundial.
21. Manejo y análisis de grandes cantidades de información.
22. Interpretar grandes cantidades de datos.
23. Una base de datos.
24. Recogida de datos a gran nivel, algunos no cuantificables de forma habitual.
25. Empleo de inteligencia artificial o redes neuronales para el manejo de grandes cantidades de datos con el objetivo de encontrar relaciones, señales, patrones... y poder hacer predicciones.
26. Conjunto de gran tamaño de datos informatizados cuyo análisis es complejo, pero aporta información útil.
27. La posibilidad de manejar gran cantidad de datos de manera muy rápida y accesible que provienen de numerosas fuentes fiables.
28. Adquirir conocimientos mediante el análisis de muchos datos.
29. Conjunto de datos procedentes de diversas bases de información.
30. Enorme cantidad de datos producidos en el día a día, procesados mediante procesos informáticos para obtener información a gran escala.
31. Una recopilación de datos a partir de medios digitales para su explotación de cara a identificar tendencias.
32. El estudio de datos a gran escala, volumen poblacional y con casi infinitas variables.
33. Grandes cantidades de datos que pueden orientar al analizarlos sobre diferentes características, tendencias o necesidades del grupo poblacional.
34. Conjunto de datos procedentes de diversas fuentes.
35. Interpretación e incluso toma de decisiones a partir de una masa de datos recopiladas mediante inteligencia artificial.
36. Datos masivos.

37. Análisis matemáticos de bases de datos con un gran número de casos y variables.
38. Manejo de un gran número de datos.
39. Herramienta informática que reúne datos de pacientes para estimar la probabilidad de que desarrollen un evento.
40. Grupo de muchos datos con muchas variables.
41. Grandes bases de datos.
42. Un conjunto de instrumentos y funciones que permiten el análisis y manejo de grandes cantidades de datos.
43. Muchos datos.
44. Gran cantidad de datos cuya gestión y análisis es difícil sin ayuda informática.

Conclusiones: A día de hoy, hemos podido observar que hay un gran nivel de desconocimiento sobre lo que es *Big Data*. Este dato choca con el hecho de que la gran mayoría de encuestados (98,1%) afirman que están o podrían estar interesados en formarse sobre *Big Data*, aunque solo un 18,9% haya recibido algún tipo de formación actualmente.

Bibliografía

1. Favaretto M, De Clercq E, Schneble CO, Elger BS, What is your definition of *Big Data*? Researchers' understanding of the phenomenon of the decade. PLoS One. 2020;15(2):e0228987.
2. Commission TE. *Big data* - Shaping Europe's digital future - European Union. 7 June 2022. [cited 2022 Jun 27].

0014-2565 / © 2023, Elsevier España S.L.U. y Sociedad Española de Medicina Interna (SEMI). Todos los derechos reservados.