

Introducción

Los datos y la necesidad de cuantificar desde una perspectiva personal

S.W.J. Reid

Centro Colaborador de la Organización Mundial de Sanidad Animal en Análisis de riesgos y modelización, c/o Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Hertfordshire, AL9 7TA, Reino Unido

Correo electrónico: principal@rvc.ac.uk

Palabras clave

Control de enfermedades – Cuantificación – Macrodatos – Medicina científicamente fundamentada – Modelización.

«Como digo a menudo, cuando uno puede medir aquello de lo que está hablando y puede expresarlo en cifras, ya sabe algo al respecto; pero cuando no puede medirlo, cuando no puede expresarlo numéricamente, el conocimiento que tiene de ello es más bien exiguo e insatisfactorio.» Lord Kelvin (1824–1907)

Cabría postular que la única lengua realmente unificadora de nuestro mundo moderno es la de los unos y ceros de nuestro ubicuo universo digital, justamente el ecosistema en el que se sitúa el concepto de macrodatos. Con todo, algo que no deja de ser central para que los datos que ahora obtenemos por miríadas de terabytes nos resulten útiles es nuestra aptitud para analizar, evaluar y jerarquizar y, tan importante o más, nuestra necesidad de cuantificar.

Han pasado casi 30 años desde que Hiatt y Goldman [1], profesores en la Facultad de Medicina de Harvard, publicaran en la revista *Nature* un artículo de opinión titulado «Hacer más científica la medicina» en el que refutaban la extendida idea de que, dado el creciente número de médicos que, además de su formación en medicina, contaban con formación básica en biología, había llegado la era de la medicina

científica. Hiatt y Goldman cuestionaban esta visión de las cosas, razonando que en medicina, si bien se habían integrado ciertos elementos básicos de la biología y las ciencias moleculares, las decisiones seguían reposando en apreciaciones subjetivas y se dejaban de lado, por lo esencial, varias disciplinas que podían sustentar la evaluación clínica, como la estadística, la epidemiología, el análisis de decisiones, la economía, el análisis de la relación costo-eficacia o la ciencia de los datos. Esta omisión, postulaban, tenía consecuencias en dos terrenos: en primer lugar, las investigaciones y estudios de población en materia de diagnóstico y resultados o efectos terapéuticos eran básicamente periféricos en el mundo de la investigación médica universitaria; en segundo lugar, las consideraciones cuantitativas, éticas y de eficacia en relación con el paciente tomado individualmente eran aspectos acaso incipientes, pero aún desatendidos.

El año siguiente, Wain-Hobson [2] publicaron en la misma revista un artículo dedicado a los avances de la investigación sobre virus de inmunodeficiencia humana en el que explicaban que los hallazgos que describían eran fruto de «una labor en equipo de biólogos, en este caso inmunólogos, con matemáticos». La semana siguiente, en su editorial, *Nature* retomaba esa observación, pero señalaba que la mención específica a la falta de «visión cuantitativa» de biólogos moleculares y virólogos apuntaba a una carencia mucho más amplia.

Tal vez resulte difícil de creer que, a finales del siglo XX, una de las publicaciones científicas más destacadas del mundo estuviera alertando a los círculos científicos de que tanto la investigación médica fundamental como la medicina clínica presentaban deficiencias en el terreno cuantitativo. Por lo que respecta al control de enfermedades, un ámbito de trabajo en el que cabría distinguir tres niveles de estudio, a saber, el molecular, el del paciente y el de las poblaciones, aplicando la vara de medir de la citada fórmula de Kelvin se observaban deficiencias en dos de esos tres niveles.

A escala poblacional, tanto en salud pública como en epidemiología ha habido siempre un sólido conocimiento y manejo de las cifras, por lo menos descriptivamente, y los conceptos de la epidemiología moderna

gozan de un largo y distinguido historial. Aunque la determinación de factores de riesgo de enfermedad y de correlaciones estadísticas entre una patología y la exposición a esos factores eran procedimientos muy extendidos en el ámbito de la medicina, los métodos de modelización matemática, utilizados desde hacía tantos años, eran patrimonio exclusivo de las disciplinas no clínicas, en especial la biología de micro y macroparásitos, la ecología y la biología evolutiva. Sin olvidar la filiación, rara vez reconocida, que se remonta a los trabajos de Bailey [3], cabe atribuir a un zoólogo y un físico, Anderson y May [4], la responsabilidad de la incorporación generalizada a la medicina y la veterinaria de los métodos de modelización de enfermedades infecciosas.

Vistas las cosas desde una escuela diferente, pero no menos distinguida, el epidemiólogo moderno citaría sin dudarlo a Popper [5] y a Rothman *et al.* [6], entre otros autores, como los artífices de la evolución y la aplicación de métodos cuantitativos. En esta escuela, la inferencia causal, el enfoque estadístico y el protagonismo del sesgo, así como los factores de confusión y el riesgo han sentado las bases de investigaciones cruciales tanto en medicina como en veterinaria.

En el otro extremo del espectro –la repercusión de los datos en las pruebas científicas que afectan al individuo–, la mencionada observación publicada en *Nature* fue eclipsada por un artículo escrito unos dos años antes por Guyatt *et al.* [7]. La necesidad de abordar la evaluación de prácticas óptimas aplicando un planteamiento estructurado condujo a la instauración de la moderna medicina científicamente fundamentada (o «medicina científica»), aun sabiendo que los principios subyacentes son tan antiguos como la propia medicina. Ahora bien, ¿cuáles son los datos probatorios? ¿Cuán frecuentes son las enfermedades que tratamos, cuán buenas son nuestras técnicas de diagnóstico, cuán adecuados son nuestros tratamientos, cuán eficaces son nuestras intervenciones? Estos interrogantes caben perfectamente en un texto de epidemiología y está claro que la praxis de la «medicina científica» reposa en las teorías y principios de la epidemiología y la bioestadística pero también, sobre todo, que la necesidad de cuantificar se erige de nuevo en un aspecto central. Sackett

et al. [8] propusieron una jerarquización, posteriormente modificada por Yusuf *et al.* [9], de las distintas clases de datos probatorios que se requieren para abordar en cuatro pasos la praxis de una «medicina científica» [10]:

- 1) análisis sistemáticos de múltiples ensayos aleatorizados, ciegos y controlados con placebo concebidos para estudiar cuestiones clínicas específicas
- 2) ensayos clínicos no aleatorizados con empleo de grupos comparativos históricos
- 3) series de casos no controlados (estudios no comparativos)
- 4) opiniones de especialistas y/o pruebas extrapoladas a partir de investigaciones publicadas.

Aunque esta jerarquización no goza de unánime aceptación, es importante destacar que la existencia de datos en abundancia y de técnicas cuantitativas son aspectos centrales para cada uno de estos tipos de prueba científica. El mundo veterinario ha seguido el mismo camino [11].

Alimentar la reflexión desde diversos ángulos

Veinte años después, disponemos de datos en abundancia y de una plétora de aplicaciones de los macrodatos en medicina y veterinaria [12, 13, 14, 15], con áreas de investigación que cubren todo el espectro que va de los datos meteorológicos y espaciales a los generados por secuenciación de genomas completos y técnicas de bioinformática. El conjunto de estudios y reflexiones reunidos en este número de la *Revista científica y técnica* de la Organización Mundial de Sanidad Animal (OMSA), impulsado por la creación del Centro Colaborador de la OMSA en Análisis de riesgos y modelización (iniciativa conjunta del Royal Veterinary College y la Agencia de Sanidad Animal y Vegetal del Reino Unido), ilustra el cuándo, el cómo y el porqué de la utilización de datos y de la aplicación de técnicas cuantitativas. Los estudios aquí descritos cubren desde la escala macro (poblaciones) hasta la microbiológica y la molecular y, en cada nivel de resolución, la cuestión

de los datos y la aplicación de un método cuantitativo son elementos fundamentales para la solidez del discurso científico y del proceso de inferencia conexo.

En este número se tocan varios temas de importancia, desde aspectos relacionados con la gobernanza hasta cuestiones referidas a una especie en concreto, pasando por ejemplos de diversos ámbitos y, algo muy importante, la utilidad y el liderazgo de la OMSA en temas fundamentales que tienen que ver con el uso de los datos con fines de vigilancia, control de enfermedades y trabajo normativo. Tenemos mucho que aprender comparando, contrastando y cotejando estas numerosas contribuciones, cuyos ámbitos de estudio van desde el sector de la producción animal para el consumo humano hasta los programas sobre animales salvajes o de compañía, desde la secuenciación de genomas completos hasta los sistemas de información geográfica, desde la estadística y las matemáticas hasta las disciplinas científicas que se ocupan de la información y los datos.

La auténtica belleza de la propuesta de este número es que aporta el ingrediente esencial para que una investigación llegue a traducirse en un relato coherente: si en todos los estudios aquí publicados la motivación científica aporta la trama y las observaciones biológicas son los verbos y sustantivos, los elementos cuantitativos constituyen los adverbios, adjetivos, pronombres y matices gramaticales que resultan en una prosa esmerada y comprensible.

Todos los estudios aquí presentados ejemplifican el lugar central de los datos y la cuantificación en el ámbito de la sanidad animal, y, aunque la bibliografía abunda en máximas «cuantitativas» menos que elogiosas, desde la que afirma que «todos los modelos son erróneos, pero algunos son útiles» [16] hasta la que observa que «los errores resultantes de datos inadecuados son menores que los resultantes de la falta de datos» (Babbage, 1792–1871), pasando por aquella sentencia (atribuida a Mark Twain y a Benjamin Disraeli) que habla de «mentiras, grandes mentiras y estadísticas», la realidad es que, aunque se pueden hacer inferencias falsas o defectuosas a partir de datos empleando técnicas cuantitativas [17], resulta mucho más fácil extraer

conclusiones inapropiadas cuando no se utilizan estas cuantificaciones o cuando no se dispone de datos.

Lo que queda por delante

La siguiente dificultad que aguarda al sector será la de abordar la síntesis de datos trabajando verdaderamente en el contexto de «Una sola salud», o dicho de otro modo: ¿cómo asegurarnos, desde el sector de la sanidad animal, de que nuestros datos puedan combinarse eficazmente con los de otros sectores u organismos, como la Organización de las Naciones Unidas para la Alimentación y la Agricultura o la Organización Mundial de la Salud [18], que habrán sido recogidos y estudiados con otros fines, sabiendo que la utilidad y las sinergias residen en la posibilidad de intersección de los conjuntos de datos? Un ejemplo lo tenemos en la lucha contra las resistencias a los antimicrobianos [19] y en la necesidad de utilizar múltiples fuentes de datos para lograr progresos reales y promover procesos decisorios que partan de datos factuales y reposen en consideraciones científicas. Este, sin embargo, es otro cantar y será objeto de otro número de la *Revista de la OMSA*.

Tras desplegar este nutrido argumentario en favor de los datos y la cuantificación, se impone para acabar una puntualización importante, puesto que ni los unos ni la otra, aun siendo necesarios, bastan por sí solos, y un planteamiento reduccionista no resultaría menos aberrante a la hora de vehicular un mensaje científico complejo. Las palabras también importan [20], como es de esperar que evidencien los artículos aquí reunidos.

Referencias

[1] Hiatt H. y Goldman L. (1994). – Making medicine more scientific. *Nature*, 371 (6493), 100. <https://doi.org/10.1038/371100a0>

- [2] Wain-Hobson S. (1995). – Virological mayhem. *Nature*, **373** (6510), 102. <https://doi.org/10.1038/373102a0>
- [3] Bailey N.T.J. (1975). – The mathematical theory of infectious diseases and its application. 2^a ed. Griffin, Londres, Reino Unido, 413 págs.
- [4] Anderson R.M. y May R.M. (1992). – Infectious diseases of humans: dynamics and control. Oxford University Press, Oxford, Reino Unido, 766 págs.
- [5] Popper K.R. (1935). – Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft. Traducción en inglés: The logic of scientific discovery. Routledge Classics, Londres, Reino Unido, 1959, 513 págs. Disponible en: <http://philotextes.info/spip/IMG/pdf/popper-logic-scientific-discovery.pdf> (consultado el 13 de marzo de 2023). Traducción en español: La lógica de la investigación científica (V. Sánchez de Zavala, tr.), Editorial Tecnos, Madrid, España, 1962, 456 págs.
- [6] Rothman K.J., Greenland S. y Lash T.L. (2008). – Modern epidemiology. 3^a ed. Lippincott Williams and Wilkins, Filadelfia, Estados Unidos de América, 758 págs. Disponible en: http://students.aiu.edu/submissions/profiles/resources/onlineBook/a9c7D5_Modern_Epidemiology_3.pdf (consultado el 13 de marzo de 2023).
- [7] Guyatt G.H. y Grupo de Trabajo sobre Medicina Basada en la Evidencia (1992). – Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*, **268** (17), 2420–2425. <https://doi.org/10.1001/jama.1992.03490170092032>
- [8] Sackett D.L., Haynes R.B., Guyatt G.H. y Tugwell P. (1991). – Clinical epidemiology: a basic science for clinical medicine. 2^a ed. Lippincott Williams and Wilkins, Filadelfia, Estados Unidos de América, 466 págs.

- [9] Yusuf S., Cairns J.A., Camm A.J., Fallen E.L. y Gersh B.J. (2008). – Evidence-based cardiology. 2^a ed. BMJ Publishing Group, Londres, Reino Unido, 1024 págs.
- [10] Sackett D.L., Straus S.E., Richardson W.S., Rosenberg W. y Haynes R.B. (2000). – Evidence-based medicine: how to practice and teach EBM. 2^a ed. Churchill Livingstone, Toronto, Canadá, 280 págs.
- [11] Cockcroft P.D. y Holmes M.A. (2003). – Handbook of evidence-based veterinary medicine. Blackwell Publishing, Oxford, Reino Unido, 210 págs. <https://doi.org/10.1002/9780470690833>
- [12] Ristevski B. y Chen M. (2018). – Big data analytics in medicine and healthcare. *J. Integr. Bioinform.*, **15** (3), 20170030. <https://doi.org/10.1515/jib-2017-0030>
- [13] Kao R.R., Haydon D.T., Lycett S.J. y Murcia P.R. (2014). – Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends Microbiol.*, **22** (5), 282–291. <https://doi.org/10.1016/j.tim.2014.02.011>
- [14] Ouyang Z., Sargeant J., Thomas A., Wycherley K., Ma R., Esmailbeigi R., Versluis A., Stacey D., Stone E., Poljak Z. y Bernardo T.M. (2019). – A scoping review of ‘big data’, ‘informatics’, and ‘bioinformatics’ in the animal health and veterinary medical literature. *Anim. Health Res. Rev.*, **20** (1), 1–18. <https://doi.org/10.1017/S1466252319000136>
- [15] Gulyaeva M., Huettmann F., Shestopalov A., Okamoto M., Matsuno K., Chu D.-H., Sakoda Y., Glushchenko A., Milton E. y Bortz E. (2020). – Data mining and model-predicting a global disease reservoir for low-pathogenic Avian Influenza (AI) in the wider pacific rim using big data sets. *Sci. Rep.*, **10** (1), 16817. <https://doi.org/10.1038/s41598-020-73664-2>

[16] Box G.E.P. (1979). – Robustness in the strategy of scientific model building. *En* Robustness in statistics (R.L. Launer y G.N. Wilkinson, coords). Academic Press, Nueva York, Estados Unidos de América, 201–236. <https://doi.org/10.1016/B978-0-12-438150-6.50018-2>

[17] Ioannidis J.P.A. (2005). – Why most published research findings are false. *PLoS Med.*, 2 (8), e124. <https://doi.org/10.1371/journal.pmed.0020124>

[18] Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), Organización Mundial de Sanidad Animal (OMSA) y Organización Mundial de la Salud (OMS) (2017). – The Tripartite’s commitment: providing multi-sectoral, collaborative leadership in addressing health challenges. FAO, OMSA & OMS, Roma, Italia, 4 págs. Disponible en: <https://doc.woah.org/dyn/portal/index.xhtml?page=alo&aloId=34824&spaceId=100> (consultado el 13 de marzo de 2023).

[19] Organización Mundial de la Salud (OMS), Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), Organización Mundial de Sanidad Animal (OMSA) y Programa de las Naciones Unidas para el Medio Ambiente (PNUMA) (2022). – Marco Estratégico de Colaboración sobre la Resistencia a los Antimicrobianos – juntos por «Una sola salud». OMS, FAO, OMSA y PNUMA, Ginebra, Suiza, 19 págs. Disponible en: <https://www.who.int/es/publications/i/item/9789240045408> (consultado el 13 de marzo de 2023).

[20] Boyle D. (2002). – The tyranny of numbers: why counting can’t make us happy. Harper Perennial GB, Londres, Reino Unido, 256 págs.

© 2023 Reid S.W.J.; licensee the World Organisation for Animal Health. This is an open access article distributed under the terms of the Creative Commons Attribution IGO Licence (<https://creativecommons.org/licenses/by/3.0/igo/legalcode>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. In any reproduction of this article there should not be any suggestion that WOAHA or this article endorses any specific organisation, product or service. The use of the WOAHA logo is not permitted. This notice should be preserved along with the article's original URL.

Pre-print