PSA PHILOSOPHY OF SCIENCE ASSOCIATION   PHILOSOPHY OF SCIENCE

SYMPOSIA PAPER

Is Data Science Transforming Biomedical Research? Evidence, Expertise, and Experiments in COVID-19 Science

Sabina Leonelli

Exeter Centre for the Study of the Life Sciences (Egenis), University of Exeter, Exeter, UK
Email: s.leonelli@exeter.ac.uk

Abstract

Biomedical deployments of data science capitalize on vast, heterogeneous data sources. This promotes a diversified understanding of what counts as evidence for health-related interventions, beyond the strictures associated with evidence-based medicine. Focusing on COVID-19 transmission and prevention research, I consider the epistemic implications of this diversification of evidence in relation to (1) experimental design, especially the revival of natural experiments as sources of reliable epidemiological knowledge; and (2) modeling practices, particularly the recognition of transdisciplinary expertise as crucial to developing and interpreting data models. Acknowledging such shifts in evidential, experimental, and modeling practices helps avoid harmful applications of data-intensive methods.

1. Introduction

Data science, and related data infrastructures and analytic tools, are frequently invoked as a major factor underpinning contemporary transformations in medical research, diagnosis, and treatment. This article considers the impact of data science on biomedical research, focusing on implications for experimental design, modeling strategies, and evidential standards, and taking the first two years of research on the COVID-19 pandemic as a case study.

I start with a sketch of current debates around biomedical evidence, pointing to the opportunities offered by data science to capitalize on vast and heterogeneous data sources, the related shift away from the highly regimented approach to data production championed by the evidence-based medicine (EBM) movement, and the emergence of a more diversified understanding of what may count as empirical insight for health-related research. I then turn to the implications of this shift for early-response research on the COVID-19 pandemic, which has been under enormous pressure to generate knowledge about the SARS-COV-2 virus that may prevent harmful effects on humans, whether by developing vaccines or by limiting

transmission. I consider two areas of COVID-19 biomedicine that proved critical to the pandemic response: the development of experimental approaches to test vaccine effectiveness, which included so-called natural experiments grounded on observations collected from populations in real time, and the use of diverse forms of expertise to develop and interpret models of COVID-19 transmission patterns and effects across patient groups.

In both cases, I show that researchers made initial headway through rapid analysis of relatively homogeneous data extracted from hospital records, tracing programs, and vaccine trials, resulting in the identification of general trends at the national level. These efforts, however, ran into trouble as soon as more granular results were needed, for instance to understand the differential impact of the pandemic across neighbourhoods and patient groups, or adjudicate divergent results coming from different research groups and approaches. These problems were offset through modifications to experimental design and modeling practices, which enabled researchers to benefit from the large volume and variety of data generated as the pandemic exploded, including the observations of relevant nonscientists such as patients and their families, frontline medical staff, social services, and public health authorities.

From consideration of these examples, I argue that data science is indeed having a transformative effect on biomedical research, fostering significant changes in the evaluation of evidence, experimental methods, and modeling practices. These changes are not, however, an unavoidable consequence of introducing new technologies and data sources. Data science tools can also be deployed as a mere complement to existing research methods, thereby yielding short-term outcomes without necessarily challenging established ways of doing in biomedicine. Using data science in this way does not take full advantage of its potential to foster robust and comprehensive investigations grounded on a wider evidence base; and it involves epistemic risks because these approaches do not support fine-grained forms of contextualization and validation.

I conclude that for data science to improve the pace, effectiveness, and reliability of biomedical research in the long term, it needs to be accompanied by epistemic shifts in evidential, experimental, and modeling standards, and that such changes need to be explicitly acknowledged and supported by research institutions. This will help to prevent harmful or inappropriate applications of data science tools within biomedicine.

2. Evidence rankings and the contemporary health data ecosystem

The emergence of evidence-based medicine in the 1990s introduced a hierarchical understanding of biomedical evidence, within which different types of data are ranked as more or less reliable depending on the methods used to generate them. Observational data (including case reports and expert opinion) sit at the bottom of the ranking, while the outcomes of randomized controlled trials (RCTs) and related systematic reviews are hailed as the "gold standard" for high-quality, robust evidence (Timmermans and Berg 2003). Many philosophers have critiqued this scheme and particularly the underlying assumption that randomization ensures the statistical significance and validity of the results (Worrall 2002, 2007; Cartwright 2007, 2011),

as well as this system's disregard for mechanistic knowledge (Russo and Williamson 2007) and mistrust of experiential knowledge by doctors, patients, and their communities (Solomon 2015). A less frequently discussed implication of this approach has been the institutionalized separation of data sources and related communities of practice from each other. Data coming from animal research, clinical trials, administrative sources, and patients records have been kept in distinct silos: They are stored in data infrastructures financed by different organizations, utilizing different standards and responding to different systems of amalgamation, resulting in little if any interoperability across. The emphasis on RCT data over all others has taken pressure off attempts to link these data to other sources of relevant evidence, resulting in ever-increasing trouble with sharing and integrating data beyond specified and highly contained environments (Leonelli 2017; Fleming et al. 2017). A direct consequence of these practices and governance model is that data analysis has been largely confined within specific methodological traditions, with modeling and inferential reasoning typically applied to homogenous data of the same type, rather than bringing together data of diverse provenance, formats, and representational power.

Fields such as epidemiology and public health, whose strong interest in the social determinants of health is badly suited to RCT evidence, never stopped pushing for a more inclusive and diversified evidence base than that sanctioned by EBM. Over the last decade, these efforts received a significant boost from the emergence of widely applicable computational tools to analyze and link a large variety of data types, such as data mash-ups, Open Data systems and semantic web technology (Fleming et al. 2017). This has disrupted existing data siloes and related rankings, most obviously by expanding the boundaries of the health data ecosystem to include new sources such as social media, digitalized administrative and social services, and self-measuring devices (see Figure 1), but also through novel forms of data governance and AI-led analytics capable of modeling data in real time and across scales.

There is more to this development than the liberal approach to evidential standards long favored within some parts of epidemiology. It is a substantive shift in the types of data and analytic tools that can be put to the service of biomedical research, a shift on which epidemiologists have been quick to capitalize (Canali and Leonelli 2022). These novel forms of data and related work have opened a new front of critique against the EBM hierarchy of evidence. Traditional boundaries between research and clinical data have started to crumble, as exemplified by the status acquired by electronic health records as medical evidence (Tempini and Teira 2020); epidemiological concepts like "exposure" have become foci for interdisciplinary research, resulting in a reconceptualization of the relationship between human health and environmental stressors (Canali and Leonelli 2022); and precision medicine has brought attention to cross-sector evidence for relevant biomarkers, though with a tendency to privilege molecular data (Prainsack 2017; Tabery 2023).

3. The COVID-19 challenge: Emergency research and fluctuating evidential standards

The onset of the COVID-19 pandemic, and the related imperative to pool international efforts toward producing relevant biomedical knowledge, exemplifies how this novel
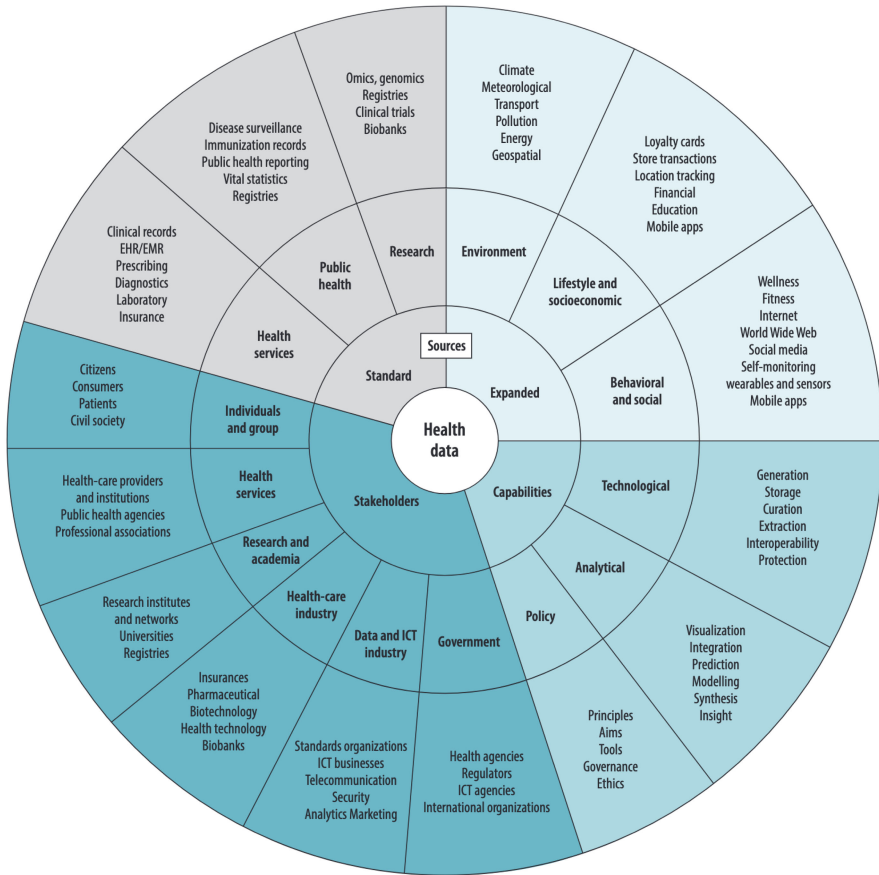
Figure 1. The health data ecosystem in 2016. *Source*: World Health Organization, CC-BY. http://www.who.int/ehealth/resources/ecosystem/en/.

evidential landscape has affected biomedical research. Researchers involved in the pandemic response were confronted with a staggering scale of data-sharing efforts, with hundreds of data infrastructures redeployed or created from scratch to collect, visualize, and model data of relevance. By February 2023, the World Health Organization's COVID-19 Database had over 800,000 entries, most of them consisting of heterogeneous and extensive datasets in their own right, and certainly not exhaustive of the myriad data initiatives in the wake of the pandemic. Key sources for patient data were hospitals and clinics, while sampling facilities around the world provided information about emerging SARS-COV-2 variants. Many nontraditional sources of health information were also recognized as research assets, including aggregated phone-derived mobility data, open government data (e.g., public use of transportation and public facilities), social media, and web mining (Zhang et al. 2021). Given the breakneck speed at which vaccines and public health measures were developed, tested, and updated, it may be argued that this enormous data sharing effort successfully fostered fast-paced research toward tackling the emergency.

This effort, however, required a shift in what was considered to be a relevant and credible evidence base for research, which came with significant challenges. Those included concerns around data access, comparability, and standardization. Acquiring data from healthcare facilities and mobile phone carriers proved expensive and not always feasible (Piasecky and Cheah 2022; Tempini 2022); data provenance was often unclear and adherence to meta-data standards was poor, when such standards were available at all (Alan Turing Institute 2021); the divide between digitalized and analogue data sources proved difficult to bridge (Ada Lovelace 2022); and existing data siloes resisted breaching (Krige and Leonelli 2021; Office for Statistics Regulation 2022). In turn, concerns were raised around the quality, representativeness, and reliability of the data, as well as the extent to which confounding factors could be accounted for (Ada Lovelace Institute 2022).

Paradoxically, this encouraged some degree of conservatism around which data sources may prove most credible, with some forms of evidence winning accolades as novel reference points for data-intensive biomedicine while others were regarded with suspicion. While it was widely agreed that RCTs would provide only part of the required evidence, data coming from controlled environments such as laboratories, such as for instance virological studies, were privileged over data collected by doctors and social services (Leonelli 2021); and social scientific expertise, including observational and ethnographic studies, was often dismissed in favour of predictive modeling grounded largely on homogeneous transmission data (Lohse and Canali 2021). One reason for these trends, aside from difficulties in accessing privately held data, was the perceived tractability of the data and its amenability to specific forms of computational analysis—a factor that, while practically important, provides no epistemic ground to disregard data sources requiring more laborious processing and interpretation. A review of how data have been used to inform the pandemic response highlighted how easily disseminated and digestible data visualizations were systematically privileged over complex disaggregated data sources, irrespectively of the degree of relevance and robustness of the information therein provided (Ada Lovelace Institute 2022).

In the following sections I briefly consider two cases in which such challenges emerged and note how the involvement of transdisciplinary expertise beyond artificial intelligence–enabled data mining fostered a more balanced and comprehensive evidence base.

4. From controlled to natural experiments: Investigating vaccine effectiveness

Quasiexperimental methods in epidemiology, also referred to as "natural experiments," are well-equipped to take advantage of the shifting health data ecosystem. A 2012 review of natural experiments undertaken by UK funding bodies defines them as an experimental situation in which "exposure to the event or intervention of interest has not been manipulated by the researcher" (Craig et al. 2012, 1182). Indeed, "the intervention is not undertaken for the purposes of research" (ibid., 1182) because it typically emerges in relation to sociopolitical or environmental changes that are outside the control of researchers: "[W]hereas in experimental designs, the participants are actively assigned to either the intervention or control group, quasi-experimental methods take advantage of exogenous sources of assignment to

the intervention" (Bernal et al. 2019, 1769). At the same time, "the variation in exposure and outcomes is analysed using methods that attempt to make causal inferences," thereby identifying characteristics of the naturally occurring event that can be used as variables and controls (Craig et al. 2012, 1182).

Natural experiments have long been employed to research the *effectiveness* of vaccines in preventing illness without harmful side-effects (Bernal et al. 2018), a focus that underscores the distance from the strict notion of vaccine *efficacy* typically associated to RCTs. In the words of leading epidemiologists, "efficacy trials (explanatory trials) determine whether an intervention produces the expected result under ideal circumstances. Effectiveness trials (pragmatic trials) measure the degree of beneficial effect under 'real world' clinical settings" (Gartlehner et al. 2006, 1). In her analysis of admissible evidence sources for health-related decision making, Cartwright is careful to note the dangers of this approach, but also the extent to which recourse to a broader evidence base may help mitigate these dangers: "[E]ffectiveness predictions are always dicey. Use of scientific evidence makes them far less so" (2011, 1401).

Given the availability of so many diverse data sources, it should come as no surprise that natural experiments proved fruitful for studies of the possible effects of COVID-19 vaccines. The adoption of these methods enabled researchers to capitalize on existing data on COVID-19 vaccination and infection rates, as well as the various ways in which existing data-sharing mechanisms (such as genomic databases and trusted research environments) were repurposed to inform small scale, nonclinical studies in several locations around the world, while underpinning the setup of large-scale clinical trials (Zhang et al. 2021; Leonelli 2021). Moreover, artificial intelligence applications fostered rapid data mining across spatial and temporal scales, thus maximizing the fruitfulness of new forms of evidence. The resulting studies enabled the study of populations in real time, thus helping to close the frustrating and dangerous gap typically charactering data collection and data analysis in this domain.

This was particularly effective in countries equipped with extensive and responsive data infrastructures. OpenSAFELY, a UK database collecting National Health Service (NHS) patient records, was used as a source of dynamic data about infections and vaccine coverage (Curtis et al. 2021; Chafetz et al. 2022); while in Brazil, the existence of detailed and well-curated administrative databases fostered studies of vaccine effectiveness across different parts of the population (Pescarini et al. 2021). However, the existence of well-maintained databases was not enough to ensure evidential robustness: Such sources were still far from comprehensive, and the fact that they only stored data for specific parts of the population (e.g., those with access to regular healthcare and/or digital medical services) generated potentially harmful bias. For example, a recent systematic review of effectiveness studies using natural experiments noted the lack of balance among available sources:

> [T]he most common study type is retrospective cohort study, often employing immunisation registries and medical databases. Only five studies considered asymptomatic infection among patients under investigation, frontline workers and randomly selected individuals in the community. Most cohort studies were conducted among healthcare workers undergoing routine RT-PCR testing as part of the hospital surveillance system. (Teerawattanon et al. 2022, 2)

When, as in this case, data collection happens largely under controlled hospital conditions, it fails to capture populations outside those environments. These issues become magnified in low- and middle-income countries (LMICs):

> most vaccine effectiveness studies to date have been conducted in high income countries with access to reliable and interlinked databases for COVID-19 vaccination, diagnosis and treatment. Such databases often do not exist in LMICs, meaning that countries will be employing prospective study designs, requiring a priori calculation of sample size and a clear plan to manage and report on confounders and missing data. (Teerawattanon et al. 2022, 26)

A crucial way out of such troubles is complementing data mining from existing large databases with studies by researchers specialized in the population at hand, including qualitative evaluations, observational approaches, and appropriately chosen proxies to make up for missing data, and extensive consultations with representatives of the population in question (as done by research on vaccine effectiveness within Brazilian indigenous populations; Pescarini et al. 2021, 2023). Such transdisciplinary methods provide a necessary counterpoint to decontextualized data mining and play a key role in calibrating the results to guarantee scientific reliability, robustness, and fairness. Ideally, the significance of qualitative studies and transdisciplinary consultation needs to be recognized from the outset of research and included in study design, so that the mining of secondary data employed in natural experiments is developed through appropriate understanding of the populations at hand. In the absence of such recognition, there is a substantive risk of using data science to producing studies grounded on partial evidence, whose results may benefit richer parts of the population while taking no account of—and potentially harming—less affluent and more vulnerable subjects.

5. Expanding biomedical expertise: Transdisciplinary input in COVID-19 transmission models

Another example is the production and interpretation of COVID-19 transmission models. Predictive models of COVID-19 transmission were heavily used from the very start of the pandemic to inform strategies around public health responses, especially social distancing rules, masking, and mobility restrictions. A well-known case are the models of the contagion curve developed by Imperial College London in early 2020, which were deployed to support lockdowns in the United Kingdom and United States. While these models are meant to produce actionable predictions from a wide variety of heterogeneous data (Fuller 2020), some datasets ended up being prioritized as evidence due to their tractability. The results of COVID-19 tests, for instance, were easy to obtain in a digital form and widely viewed as essential parameters for epidemic models such as SIR (susceptible-infectious-removed). By contrast, data on which hospital patients was being intubated to support respiratory function were intractable due to the great variation in intubation methods, duration, and records, which meant different hospitals were recording that information in different and often incompatible ways (Alan Turing Institute 2021; Office for Statistics Regulation 2022). The urgency of modeling as fast as possible, combined with difficulties in fitting

some datasets to the models, resulted in an evidential grounding of predictive models that was much less comprehensive than hoped for, with potentially dire consequences for the validity of the models (Leslie et al. 2021). In addition, there was the difficulty in assessing the reliability of the data that were in fact used: Test data can be uneven in the extent and manner in which they are obtained, depending on the scale and targets of testing in each country, which makes a big difference at scale. Data on pandemic deaths also proved hard to validate due to the diversity of measures used across regions, including differences in who counts as "dead" and how an association with COVID-19 was determined (Nature 2023).

Given these issues, it could be argued that predictive modeling around disease dynamics is best positioned to support qualitative conclusions (e.g., the relative efficacy of proposed interventions within highly well-specified conditions) rather than quantitative predictions (e.g., the number of people in various states at time t).[1] The results of predictive modeling thus need to be understood and contextualized through reference to other forms of expert input (Goldstein et al. 2020), and particularly forms of evidence that can document the broader socioeconomic setting within which predictions are supposed to apply (Cousins et al. 2020). This requires a reframing of the way in which such models may be said to be data "driven": The question is not how many datasets may be used to inform the models, but rather how diverse and well-curated such data are and how models should be calibrated to ensure that the modeling outputs adequately reflects the empirical input. As Frisch and colleagues have pointed out, models need to be evaluated for their performativity rather than their accuracy (van Basshuysen et al. 2021), which involves integrating quantitative measurements with qualitative observations, and paying more attention to local scenarios than to overarching trends. Key to such evaluation is appeal to transdisciplinary insights to improve COVID-19 transmission studies so that they consider social determinants of health.

Transdisciplinary input does not mean "anything goes." Rather, it involves the painstaking work of identifying and engaging communities of stakeholders with appropriate expertise, whose composition depends on the models, scenarios, and questions at hand. Some of the best early predictions on the impact of COVID-19 on human health came from models produced in consultation with existing transdisciplinary networks, such as those focused on documenting and treating specific diseases. The EULAR COVID-19 Database, for instance, was born of existing strong ties among European researchers, patient groups, doctors, and industry interested in rheumatic conditions—a community that was built over many years and could be easily and swiftly recruited in 2020 to help evaluate data and calibrate models around the risk and severity of COVID-19 for rheumatology patients (Alan Turing Institute 2021). Similarly, trusted data repositories such as the Secure Anonymised Information Linkage database in Wales, which comprises expert curators with 15+ years of experience in managing complex health data and supporting question-oriented studies with specific communities of participants, played a crucial role in the analysis of the differential impact of the pandemic on minority ethnic groups in England (ibid.).

---

[1] Similar arguments have been made in relation to predictive modeling in EBM (Cartwright 2011; Fuller and Flores 2015).

6. Conclusion

There is no doubt that the emergence of methods to collect and analyze vast and heterogenous data sources is changing biomedicine, and particularly the ways in which evidence, experiments, and models are understood and used for discovery. The change is most conspicuous when compared to the EBM canon, according to which randomized controlled trials constitute a gold standard for evidence while data coming from other sources, and particularly observational data, are conceptualized as suspicious and unreliable. Through some examples of recent, data-intensive research on COVID-19, I have argued that the rise of data science as a crucial component of biomedicine is helping to promote a broader understanding of evidence, and that this has implications for experimental design as well as modeling practices.

I conclude that such changes are crucial to the deployment of novel methods and instruments for data mining and modeling, and thus to the application of artificial intelligence within biomedical research. It is not simply the deployment of new computational tools that marks a shift in biomedical practice, but rather the ways in which existing methods and practices are adapted to benefit from such tools. In other words, implementing a novel machine learning approach to biomedical discovery, for instance to identify biochemical compounds to lessen unpleasant symptoms or engineer gene products to treat hereditary disease, is not a matter of bringing a new toy into a lab and expecting it to substitute relevant human expertise, but rather of restructuring the research design to ensure the new insights are appropriately supported and evaluated.

The preceding examples indicate how data produced by patient associations, qualitative researchers, medical doctors, and frontline hospital staff proved fundamental to the pandemic response not solely by virtue of being collected and analyzed on a large scale and with the help of computational tools, but by virtue of being incorporated into an understanding of experimentation that recognizes the value of real-life observations alongside data acquired under controlled conditions, and an understanding of modeling that recognizes the value of frontline experiences by doctors and patients, as well as the input of social scientists and public health experts, in calibrating and interpreting data models. Acknowledging such shifts in evidential, experimental, and modeling practices is essential to avoid harmful applications of data-intensive methods.

Data science can best inform biomedical research when it helps to address, rather than entrench, data biases; consider alternative visions of relevant interventions; utilize multiple forms of health-related expertise, some of which may emerge from research on the social determinants of health, some of which may come from outside professional science altogether; and promote mechanisms to validate and continuously verify the reliability of algorithms used to automate data analysis. A large outstanding challenge remains the role of technology companies, and particularly large corporation such as Google and Amazon, in pushing techno-determinist utopias where artificial intelligence–powered automation is privileged over human-in-the-loop approaches to biomedical research and interventions. The market imperative to save costs by choosing faster, automated solutions with little space for human feedback and input is in direct tension with the recognition of broad transdisciplinary expertise as indispensable to contextualizing data, designing studies, and calibrating models.

References

Ada Lovelace Institute. 2022. *A Knotted Pipeline: Data-Driven Systems and Inequalities in Health and Social Care.* Report. https://www.adalovelaceinstitute.org/report/knotted-pipeline-health-data-inequalities

Alan Turing Institute. 2021. *Data Science and AI in the Age of COVID-19.* Report. https://www.turing.ac.uk/research/publications/data-science-and-ai-age-covid-19-report?_cldee=cy5sZW9uZWxssaUBleGV0ZXIuY WMudWs%3D&esid=d69931de-fdc9-eb11-bacc-000d3ad5fb39&recipientid=contact-b8720eeae430e81181 1970106faa5221-dd75210dae0542ca93a70fba8f44331f

Bernal, James A Lopez, Nick Andrews, and Gayatri Amirthalingam. 2019. "The Use of Quasi-experimental Designs for Vaccine Evaluation." *Clinical Infectious Diseases* 68(10): 1769–1776. https://doi.org/10.1093/cid/ciy906

Canali, Stefano, and Sabina Leonelli. 2022. "Reframing the Environment in Data-Intensive Health Sciences." *Studies in the History and Philosophy of Science* 93:203–14.

Cartwright, Nancy. 2007. "Are RCTs the Gold Standard?" *BioSocieties* 2:11–20.

Cartwright, Nancy. 2011. "A Philosopher's View of the Long Road from RCTs to Effectiveness." *The Lancet* 377:1400–1.

Chafetz, Hannah, et al. 2022. *The #Data4COVID19 Review: Assessing the Use of Non-traditional Data during a Pandemic Crisis.* Technical Report, GovLab. https://review.data4covid19.org

Cousins, Thomas, Sabina Leonelli, Michelle Pentacost, and Kaushik Sunder Rajan. 2020. "Situating the Biology of COVID-19: A Conversation on Disease and Democracy." *The India Forum.* https://www.theindiaforum.in/article/situating-biology-covid-19

Craig, Peter, et al. 2012. "Using Natural Experiments to Evaluate Population Health Interventions: New Medical Research Council Guidance." *Journal of Epidemiology Community Health* 66:1182–86. https://doi.org/10.1136/jech-2011-200375

Curtis, Helen J., et al. 2021. "Trends and Clinical Characteristics of COVID-19 Vaccine Recipients: A Federated Analysis of 57.9 Million Patients' Primary Care Records In Situ Using OpenSAFELY." *British Journal of General Practitioners* 72 (714):e51–62.

Fleming, L. E., Tempini, N., Gordon-Brown, H., Nichols, G., Sarran, C., Vineis, P., Leonardi, G., Golding, B., Haines, A., Kessel, A., Murray, V., Depledge, M., Leonelli, S. 2017. *Big Data in Environment and Human Health: Challenges and Opportunities. Oxford Encyclopaedia for Environment and Human Health.* Oxford University Press. DOI: 10.1093/acrefore/9780199389414.013.541.

Fuller, Jonathan. 2020. "Models vs. Evidence." *Boston Review.* http://bostonreview.net/science-nature/jonathan-fuller-models-v-evidence

Fuller, Jonathan, and Luis J. Flores. 2015. "The Risk GP Model: The Standard Model of Prediction in Medicine." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 54:49–61. https://doi.org/10.1016/j.shpsc.2015.06.006

Gartlehner, Gerald, et al. 2006. *Criteria for Distinguishing Effectiveness from Efficacy Trials in Systematic Reviews.* Rockville, MD: Agency for Healthcare Research and Quality. Technical Review 12. https://www.ncbi.nlm.nih.gov/books/NBK44024/

Krige, John, and Sabina Leonelli. 2021. "Mobilizing the Translational History of Knowledge Flows: COVID-19 and the Politics of Knowledge at the Borders." *History and Technology* 37 (1):125–46.

Leonelli, Sabina. 2017. *Biomedical Knowledge Production in the Age of Big Data.* Report for the Swiss Science and Innovation Council. http://www.swir.ch/images/stories/pdf/en/Exploratory_study_2_2017_Big_Data_SSIC_EN.pdf

Leonelli, Sabina. 2021. "Data Science in Times of Pan(dem)ic." *Harvard Data Science Review* 3 (1). https://doi.org/10.1162/99608f92.fbb1bdd6

Leslie, David, et al. 2021. "Does 'AI' Stand for Augmenting Inequality in the Era of COVID-19 Healthcare?" *British Medical Journal* 372. https://www.bmj.com/content/372/bmj.n304

Lohse, Simon, and Stefano Canali. 2021. "Follow *the* Science? On the Marginal Role of the Social Sciences in the COVID-19 Pandemic." *European Journal of Philosophical Science* 11:99. https://doi.org/10.1007/s13194-021-00416-y

Nature. 2023. "Editorial: Missing Data Mean We'll Probably Never Know How Many People Died of COVID." *Nature* 612 (7940).

Office for Statistics Regulation (2022). *2022 Update: Lessons Learned for Health and Social Care Statistics from the COVID-19 Pandemic.* https://osr.statisticsauthority.gov.uk/wp-content/uploads/2022/11/Lessons_learned_for_health_and_social_care_statistics_from_the_COVID-19_pandemic_2022.pdf

Pescarini, Julia M., et al. 2021. "Methods to Evaluate COVID-19 Vaccine Effectiveness, with an Emphasis on Quasi-experimental Approaches." *Cien Saude Colet* 26 (11):5599–5614. https://doi.org/10.1590/1413-812320212611.18622021

Pescarini, Julia M., et al. 2023. "Vaccine Coverage and Effectiveness against Laboratory-Confirmed Symptomatic and Severe COVID-19 in Indigenous People in Brazil: A Cohort Study." PREPRINT (Version 1), *Research Square.* https://doi.org/10.21203/rs.3.rs-2550459/v1

Piasecki, Jan, and Phaik Yeong Cheah. 2022. "Ownership of Individual-Level Health Data, Data Sharing, and Data Governance." *BMC Medical Ethics* 23:104. https://doi.org/10.1186/s12910-022-00848-y

Pranisack, Barbara. 2017. *Personalised Medicine: Empowered Patients in the 21st Century?* New York: New York University Press.

Russo, Federica, and Jon Williamson. 2007. "Interpreting Causality in the Health Sciences." *International Studies in the Philosophy of Science* 21:157–70.

Solomon, Miriam. 2015. *Making Medical Knowledge.* Oxford: Oxford University Press.

Tabery, James. 2023. *Tyranny of the Gene: Personalised Medicine and Its Threat to Public Health.* New York, NYC: Knopf.

Teerawattananon Yot, Thunyarat Anothaisintawee, Chatkamol Pheerapanyawaranun, Siobahn Botwright, Katica Akksilp, Natchalaikorn Sirichumroonwit, et al. 2022. "A Systematic Review of Methodological Approaches for Evaluating Real-World Effectiveness of COVID-19 Vaccines: Advising Resource-Constrained Settings." *PLoS ONE* 17 (1):e0261930. https://doi.org/10.1371/journal.pone.0261930

Tempini, Niccolo. 2022. "Pandemic Data Circulation." *Technoscienza* 13 (1):71–95.

Tempini, Niccolo, and Sabina Leonelli. 2021. "Actionable Data for Precision Oncology: Framing Trustworthy Evidence for Exploratory Research and Clinical Diagnostics." *Social Science and Medicine* 272: 113760. https://doi.org/10.1016/j.socscimed.2021.113760

Tempini, Niccolo, and David Teira. 2020. "The Babel of Drugs: On the Consequences of Evidential Pluralism in Pharmaceutical Regulation and Regulatory Data Journeys." In *Data Journeys in the Sciences*, edited by Sabina Leonelli and Niccolo Tempini. Cham: Springer, pp. 207–225. https://doi.org/10.1007/978-3-030-37177-7_11

Timmermans, Stephan, and Mark Berg. 2003. *The Gold Standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care.* Philadelphia: Temple University Press.

UK Statistics Authority. 2021. *Inclusive Data Taskforce Recommendations Report: Leaving No One Behind—How Can We Be More Inclusive in Our Data?*, Section 5.

van Basshuysen, Philippe, et al. 2021. "Three Ways in Which Pandemic Models May Perform a Pandemic." *Erasmus Journal for Philosophy and Economics* 14 (1):110–27. https://doi.org/10.23941/ejpe.v14i1.582.

World Health Organization. 2016. *eHealth. The Health Data Ecosystem and Big Data.* Geneva: World Health Organization. http://www.who.int/ehealth/resources/ecosystem/en/

Worrall, John. 2002. "What Evidence in Evidence-Based Medicine." *Philosophy of Science* 69:S316–330.

Worrall, John. 2007. "Evidence in Medicine and Evidence-Based Medicine." *Philosophy Compass* 2:981–1022.

Zhang, Quinpeng, et al. 2021. "Data Science Approaches to Confronting the COVID-19 Pandemic: A Scoping Review." *Philosophical Transactions of the Royal Society: Part A* 380:20210127.