BMJ Health & Care Informatics

Data science to analyse the largest natural experiment of our time

Edward Christopher Dee ,[1] Joseph A Paguio,[2] Jasper S Yao,[2] Aaron Stupple,[3] Leo Anthony Celi [4,5]

To cite: Dee EC, Paguio JA, Yao JS, *et al*. Data science to analyse the largest natural experiment of our time. *BMJ Health Care Inform* 2020;**27**:e100177. doi:10.1136/bmjhci-2020-100177

Received 16 May 2020
Accepted 08 July 2020

[1]Harvard Medical School, Boston, Massachusetts, USA
[2]University of the Philippines Manila College of Medicine, Manila, Metro Manila, Philippines
[3]Baystate Medical Center, Springfield, Massachusetts, USA
[4]Division of Pulmonary Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
[5]Laboratory for Computational Physiology, Massachusetts Institute of Technology, Cambridge, MA, United States

Correspondence to
Dr Leo Anthony Celi;
lceli@mit.edu

Pandemics have left indelible marks on medicine. The 1918 influenza accelerated the emergence of American medicine as a scientific enterprise.[1] HIV ushered in universal precautions as the minimum requirement for the care of every patient.[2] One defining impact of COVID-19 might be the widespread adoption of telemedicine.[3] However, though telemedicine is a change out of necessity, should we look for other opportunities to transform medicine for the better?

Consequences of the unprecedented national lockdowns include massive reductions in emergency room visits and hospitalisations, as well as cancellations of the majority of elective procedures. Some decrease in medical care would be expected during such a crisis. However, the magnitude of decrease in essential care is striking: for example, cardiac catheterisation for ST-elevation myocardial infarction was reduced by 38% in some areas during peaks in infection rates.[4] Stroke presentations in some hospitals were halved during the height of the pandemic.[5] Many surgical procedures were and continue to be delayed to a safer time.[6] In some areas, endoscopic and colonoscopic procedures are only advised if absolutely necessary, weighed against the risk of contracting coronavirus.[7]

While it is clear that missing some treatments is harmful—undiagnosed or untreated heart attacks precipitate cardiac arrest, heart failure, or death—the avoidance of other treatments may be beneficial. We should take advantage of this opportunity to investigate which of the deferred interventions matter as much as we think they do. COVID-19 presents us with a natural experiment to study changes that otherwise would have been logistically impossible, seemingly unethical, or culturally incongruous. What happens when we delay particular procedures? Which procedures are more vital than others in terms of improving quality of life? Critically: how often do our tests and treatments lack clear value?

Answers to these questions may inform care decisions as nations face subsequent peaks in infection rates as well as future epidemics and pandemics. Similarly, even after the COVID-19 pandemic, studies of the consequences of omissions and delays done to minimise COVID-19 infection risk may inform ways in which systems can limit low-value care and instead focus resources on high-value care.

There is precedent for such investigations. A 2018 study in the *Journal of the American Heart Association* found that Medicare beneficiaries hospitalised at academic centres with acute myocardial infarctions experienced different outcomes based on whether or not they were hospitalised during dates of a major annual interventional cardiology meeting; a subgroup of patients who did not receive percutaneous coronary interventions (PCI) experienced improved outcomes if they were admitted during meeting days.[8] Patients who presented on meeting days were pseudo-randomised to practice patterns of providers who chose not to attend the meeting, creating a natural experiment that enabled the study of non-procedural treatment of some of these high-risk patients. Today, a similar scenario is playing out on a larger scale. Patients who are not receiving treatment because of the crisis have been pseudo-randomised to the non-treatment arm.

In light of such pseudo-randomisation due to the pandemic, statistical techniques such as causal inference analysis will allow the use of what will become retrospective data, as the future unfolds, to ask questions in ways that simulate a randomised controlled trial (RCT). The last decade has seen a surge in the use of data science and artificial intelligence in clinical research. However, many studies using machine learning techniques have focused on prediction of events, trajectories and outcomes by discovering patterns and associations, leading to models that

despite high accuracy, fail as soon as they are deployed as tools in the real world.[9] Applying causal inference approaches to retrospective analyses, whether making use of artificial intelligence techniques or less complex studies, may facilitate more generalisable research whose conclusions inform the provision of value-based care after the pandemic.

Simply discovering and modelling the relationship between features to make a prediction, classification, or optimisation without acknowledging causal pathways may lead to errors when models are applied to settings that are not identical to those from which the training and test data were obtained. Furthermore, although the gold standard for inferring causality is the RCT, performing RCTs for every question, for every patient population, for every clinical context, and for every temporal change in practice patterns is simply not feasible. The causal inference methodologies were created to leverage observational data to estimate effect size when treatment decisions are not randomised.[10] Many of these studies are cheaper and quicker than RCTs and may be used to answer questions for which an RCT would be unethical or would itself be low in value.

One approach, called inverse probability weighting (IPW), attempts to approximate the outcome if all subjects were assigned either treatment or non-treatment arm by weighting each subject's data in a manner that is inversely proportional to the likelihood of assignment to each arm. By inflating the weight of subjects who are under-represented, IPW tries to mitigate the biases due to differential assignment.[10]

A recent study by Faridi *et al* used IPW to assess the outcomes of ad hoc as opposed to delayed PCI in over half a million patients with stable coronary artery disease and found that ad hoc PCI was associated with a lower risk of bleeding but no difference in risk of kidney injury or mortality.[11] Unlike RCTs, treatment decisions in reality are far from random; effects attributed to ad hoc as opposed to delayed PCI are confounded by treatment selection bias. Therefore, weights were assigned based on the inverse probability of ad hoc versus delayed intervention, taking into consideration available demographic, procedural, clinical and hospital data. IPW allows minimisation of observed confounding and, despite the potential of unmeasured confounding, may approximate an RCT. Notably, the findings of Faridi *et al* were in line with findings from RCTs that asked comparable questions.[12 13] Other techniques also allow the study of potentially causal associations using retrospective data, such as hierarchical regression and more complex techniques such as Bayesian networks and neural networks.

Appropriate employment of causal inference techniques, which have been in existence since the 1990s, is incumbent on the availability of high-resolution data. The pandemic may grant us a large enough sample size with less patient exclusion and greater intensity of exposure over a sufficiently long observation period to conduct such analyses. Faridi *et al* made use of large amounts of high-dimensional multicentre data that allowed the authors to control for a great number of potentially confounding covariates.[11 14] We and others have used causal inference approaches with single-centre data[15]; however, high-quality data from multiple institutions may allow researchers to draw even stronger conclusions. The global nature of the current pandemic will allow even broader use of such statistical techniques in asking questions about common procedures.

Like all retrospective studies, optimising control for confounding remains a challenge. Researchers will need to identify endpoints that correspond to the particular intervention that was delayed and to tease apart outcomes that may be due to delays or deferrals in other aspects of the patient's care. The utility of causal inference techniques such as IPW will rely on the validity of assumptions and the inclusion of confounders that are based on sound clinical judgement. Careful (and critical) selection of analytical techniques should also inform studies.[16] Appropriate measures to curate the quality of the data will be necessary: quality in, quality out. It is important as well for researchers, clinicians, and health service executives to carefully define the questions whose answers may best serve their individual healthcare contexts and to assess whether or not they have appropriate data to ask these questions. Furthermore, early multidisciplinary engagement with data scientists and other specialists is critical.

This work carries weight beyond an academic exercise. In the USA, COVID-19 is yet another call to action to address the healthcare disparities that disproportionately harm minority populations.[17] One way to free resources for these groups is to identify and eliminate low-value care. Using the natural experiment that is the pandemic may help us construct a new normal that provides patients more of what they need and less of what they do not.

Twitter Edward Christopher Dee @EChrisDee and Leo Anthony Celi @MITCriticalData

Contributors LAC generated the idea and theme of the article. All authors contributed to the development of ideas herein and the drafting and finalisation of the manuscript.

Funding LAC is funded by the National Institute of Health through NIBIB R01 EB017205.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Edward Christopher Dee http://orcid.org/0000-0001-6119-0889
Leo Anthony Celi http://orcid.org/0000-0001-6712-6626

REFERENCES

1 Wooliscroft B. *The great influenza: the story of the deadliest pandemic in history*, 2008.
2 Broussard IM, Bhimji SS. *Universal precautions*, 2018.
3 Hollander JE, Carr BG. Virtually perfect? telemedicine for Covid-19. *N Engl J Med* 2020;382:1679–81.
4 Garcia S, Albaghdadi MS, Meraj PM, *et al*. Reduction in ST-segment elevation cardiac catheterization laboratory activations in

the United States during COVID-19 pandemic. *J Am Coll Cardiol* 2020;75:2871–2.
5 McFarling UL. 'Where are all our patients?': Covid phobia is keeping people with serious heart symptoms away from ERs, 2020. Available: https://www.statnews.com/2020/04/23/coronavirus-phobia-keeping-heart-patients-away-from-er/
6 Prachand VN, Milner R, Angelos P, *et al*. Medically necessary, Time-Sensitive procedures: scoring system to ethically and efficiently manage resource scarcity and provider risk during the COVID-19 pandemic. *J Am Coll Surg* 2020;231:281–8.
7 Di Saverio S, Pata F, Gallo G, *et al*. Coronavirus pandemic and colorectal surgery: practical advice based on the Italian experience. *Colorectal Dis* 2020;22:625–34.
8 Jena AB, Olenski A, Blumenthal DM, *et al*. Acute myocardial infarction mortality during dates of national interventional cardiology meetings. *J Am Heart Assoc* 2018;7. doi:10.1161/JAHA.117.008230. [Epub ahead of print: 09 Mar 2018].
9 Beede E, Baylor E, Hersch F, *et al*. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. *CHI Conf Hum Factors Comput Syst* 2020:1-–12.
10 Mansournia MA, Altman DG. Inverse probability weighting. *BMJ* 2016;352:i189.
11 Faridi KF, Rymer JA, Rao SV, *et al*. Ad hoc percutaneous coronary intervention in patients with stable coronary artery disease: a report from the National cardiovascular data registry CathPCI registry. *Am Heart J* 2019;216:A257.
12 Boden WE, O'Rourke RA, Teo KK, *et al*. Optimal medical therapy with or without PCI for stable coronary disease. *N Engl J Med* 2007;356:1503–16.
13 Stergiopoulos K, Brown DL. Initial coronary stent implantation with medical therapy vs medical therapy alone for stable coronary artery disease: meta-analysis of randomized controlled trials. *Arch Intern Med* 2012;172:312–9.
14 Harrell FE, Lee KL, Califf RM, *et al*. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984;3:143–52.
15 Yao JS, Paguio JA, Dee EC, *et al*. The minimal effect of zinc on the survival of hospitalized patients with Covid-19: an observational study. *Chest* 2020.
16 Federspiel JJ, Anstrom KJ, Xian Y, *et al*. Comparing inverse probability of treatment weighting and instrumental variable methods for the evaluation of adenosine diphosphate receptor inhibitors after percutaneous coronary intervention. *JAMA Cardiol* 2016;1:655.
17 Laurencin CT, McClinton A. The COVID-19 pandemic: a call to action to identify and address racial and ethnic disparities. *J Racial Ethn Health Disparities* 2020.