COVID-19 and the epistemology of epidemiological models at the dawn of AI

George TH Ellison
Centre for Data Innovation, Faculty of Science & Technology, University of Central Lancashire, Preston, UK;
gellison@uclan.ac.uk; ORCID: 0000-0001-8914-6812

Summary

The models used to estimate disease transmission, susceptibility and severity determine what epidemiology can (and cannot tell) us about COVID-19. These include: 'model organisms' chosen for their phylogenetic/aetiological similarities; multivariable statistical models to estimate the strength/direction of (potentially causal) *relationships* between variables (through 'causal inference'), and the (past/future) *value* of unmeasured variables (through 'classification/prediction'); and a range of modelling techniques to predict beyond the available data (through 'extrapolation'), compare different hypothetical scenarios (through 'simulation'), and estimate key features of dynamic processes (through 'projection'). Each of these models: address different questions using different techniques; involve assumptions that require careful assessment; and are vulnerable to generic and specific biases that can undermine the validity *and* interpretation of their findings. It is therefore necessary that the models used: can actually address the questions posed; *and* have been competently applied. In this regard, it is important to stress that extrapolation, simulation and projection cannot offer accurate predictions of future events when the underlying mechanisms (and the contexts involved) are poorly understood and subject to change. Given the importance of understanding such mechanisms/contexts, and the limited opportunity for experimentation during outbreaks of novel diseases, the use of multivariable statistical models to estimate the strength/direction of potentially causal relationships between two variables (and the biases incurred through their misapplication/misinterpretation) warrant particular attention. Such models must be carefully designed to address: 'selection-collider bias', 'unadjusted confounding bias' and 'inferential mediator adjustment bias' – all of which can introduce effects capable of enhancing, masking or reversing the estimated (true) causal relationship between the two variables examined.[1] Selection-collider bias occurs when these two variables independently cause a third (the 'collider'), and when this collider determines/reflects the basis for selection in the analysis. It is likely to affect all incompletely representative samples, although its effects will be most pronounced wherever selection is constrained (e.g. analyses focusing on infected/hospitalised individuals). Unadjusted confounding bias disrupts the estimated (true) causal relationship between two variables when: these share one (or more) common cause(s); *and* when the effects of these causes have *not* been adjusted for in the analyses (e.g. whenever confounders are unknown/unmeasured). Inferentially similar biases can occur when: one (or more) variable(s) (or 'mediators') fall on the causal path between the two variables examined (i.e. when such mediators are *caused by* one of the variables and are *causes of* the other); *and* when these mediators *are* adjusted for in the analysis. Such adjustment is commonplace when: mediators are mistaken for confounders; prediction models are mistakenly repurposed for causal inference; or mediator adjustment is used to estimate direct and indirect causal relationships (in a mistaken attempt at 'mediation analysis'). These three biases are central to ongoing and unresolved epistemological tensions within epidemiology. All have substantive implications for our understanding of COVID-19, and the future application of artificial intelligence to 'data-driven' modelling of similar phenomena. Nonetheless, competently applied and carefully interpreted, multivariable statistical models may yet provide sufficient insight into mechanisms and contexts to permit more accurate projections of future disease outbreaks.

---

[1] These biases, and the terminology involved, may be challenging to readers who are unfamiliar with the use of causal path diagrams (such as Directed Acyclic Graphs; DAGs) which have been instrumental in identifying the different roles that variables can play in causal processes (whether as 'exposures', 'outcomes', 'confounders', 'mediators', 'colliders', 'competing exposures' or 'consequences of the outcome') and revealing hitherto under-acknowledged sources of bias in analyses designed to support causal inference. For what we hoped might offer accessible introductions to DAGs (and how [not] to use these) please see: Ellison (2020); and Tennant et al. (2019). For more technical detail on 'collider bias', 'unadjusted confounding bias' and 'inferential mediator adjustment bias' (and its related concern, the 'Table 2 fallacy'), please refer to: Cook and Ranstam 2017; Munafò et al. (2018); Tennant et al. (2017); VanderWeele and Arah (2011); and Westreich and Greenland (2013).

"Since all models are wrong the scientist must be alert to what is importantly wrong.
It is inappropriate to be concerned about mice when there are tigers abroad."
George Box (1976)

Introduction

The dynamic nature of infectious disease poses particular challenges for epidemiological models, and none more so than when the disease concerned is caused by a newly emerging pathogen about which little (if anything) is known (Lloyd-Smith 2015). Rapidly identifying and characterising the pathogen responsible, and estimating the parameters that determine and reflect its transmission and severity, are all the more critical when its potential impact is global, significant and uncertain. Yet the natural history of any infectious pathogen rarely survives sustained contact with its human hosts (Jackson 2002), and the resulting changes in behaviour (both individual and social; Funk et al. 2009) further complicate efforts to understand the progression of the disease. As we draw towards the end of the world's initial encounter with COVID-19 – and with the benefit of a little hindsight – this Commentary aims to: examine the epistemological role(s) that epidemiological models have played during the course of the pandemic thus far; and reflect on George Box's (1976) aphorism regarding 'imperfect but useful models' at the dawn of 'Big data' modelling and artificial intelligence (AI).

Conceptualisation: 'model' organisms

Although not strictly 'models' in the (statistical) sense that George Box intended, the use of 'model organisms'[2] as theoretical constructs to characterise and predict the likely nature and progression of novel diseases has a long history in epidemiology (Fuller 2020a; Rettner 2020). In the case of SARS-CoV-2 – the virus responsible for COVID-19 – a number of model organisms with established pandemic potential appeared apposite, including: two recently emerging human coronaviruses (SARS-CoV-1 and MERS-CoV, which are phylogenetically related to SARS-CoV-2); and the family of human respiratory viruses responsible for seasonal influenza and the 2009-10 swine flu pandemic (Callaway et al. 2020; Petersen et al. 2020; Yee et al. 2020; see **Supplementary Table S1**). Parameter estimates from previous research on these model organisms helped to generate projections (e.g. Ioannidis et al. 2020) and simulate the 'reasonable best/worst-case scenarios' (e.g. SAGE 2020) on which rapid epidemiological and public health responses could be prepared and deployed; while the transmissibility of influenza (Dorigatti et al. 2020; Ferguson et al. 2020) coupled with the severity of SARS and MERS (Park et al. 2020) lent these efforts an added sense of urgency (Ioannidis 2020; Paules et al. 2020). However, as COVID-19 spread, and clinical data rapidly accumulated, the balance between 'imperfect' and 'useful' shifted amid growing realisation of the inherent (pathogen- and disease-specific; Whiting 2020; Paules et al. 2020) and extrinsic (outbreak- and context-specific; Burn-Murdoch and Giles 2020; Goldstein and Atherwood 2020) challenges facing the measurement of key parameters critical to the characterisation of this (or any novel) infectious disease (Lloyd-Smith 2015). While these challenges have inevitable consequences for the accuracy and precision of any such data, they also have a tendency to conflate 'uncertainty' (due to inaccuracy or imprecision in sampling and measurement) with 'variability' (due to inherent biological instability or variation) in the reported estimates of key disease characteristics (Fuller 2020b; Richardson and Speigelhalter 2020; see **Supplementary Table S1**).

Description, 'prediction' and causal inference: analytical statistical models

Epidemiologists are very familiar with the constraints that uncertainty and variability pose (Blower and Dowlatabadi 1994), though many adopt the view that – provided the risks of error and imprecision do not vary dramatically (or systematically); and provided there are enough data to moderate the risk of chance fluctuation – even suboptimal information can help provide insight, and can be used to generate foresight (Ashofteh and Bravo 2020; Ritchie et al. 2020; Woolf et al. 2020). It is on this basis that epidemiologists, statisticians and data analysts have – from the earliest stages of the COVID-19 pandemic – held their noses

[2] 'Model organisms' are also those selected or developed for investigation/experimentation under controlled (often laboratory-based) conditions.

and set to work on whatever data have been available to: visualise patterns and summarise variation in the data (e.g. Lescure et al. 2020; Simeone 2020); draw comparisons between different datasets (Ritchie et al. 2020); investigate associations between different characteristics and parameters (e.g. Korber et al. 2020); generate 'predictions'[3] (e.g. Qin et al. 2020); and infer causal mechanisms (e.g. Williamson et al. 2020). However, this is also the basis on which some data scientists are often content to use very large quantities of deeply flawed data (so-called 'Big data'; Mondal et al. 2020; Wang et al. 2020) – with little concern for error (and little regard for bias; Ayyoubzadeh et al. 2020; Qin et al. 2020; Sun et al. 2020) – to identify patterns and relationships that risk being artefacts of whatever extraneous (and intrinsic) factors determine the coverage, availability and information-value of the data available (Arnold et al. 2020).

Unlike the more obvious problems of measurement (which primarily affect the reliability, validity and consistency of the 'raw' data; Keogh et al. 2020; Raleigh 2020), the errors and biases that can arise from collating, summarising and analysing such data (Lash et al. 2014; Goldstein and Atherwood 2020; Shaw et al. 2020) – particularly when analysis extends beyond *description* to *comparison*, '*prediction'* and *causal inference* – are often hidden to all but the most expert eye (and even then may prove illusive, difficult or impossible to spot; e.g. Richiardi et al. 2013; Beggs et al. 2020).[4] While collating data and summarising these to describe their distribution within any given population (or sample) also depends in no small part on the denominator concerned (which itself may be poorly specified, misunderstood or simply overlooked; Reyna and Brainerd 2008); the principal analytical and inferential biases facing *comparative*, *'predictive'* and *causal* analyses stem not only from flaws in the sampling frame(s) used (Lash et al. 2009), but also from weaknesses in: the variables available and selected for consideration in the analyses; the parameterisation of the variables selected; the design of the statistical models used; and the interpretation of outputs therefrom (Arnold et al. 2020). These biases can compound errors and inconsistencies in measurement, making even the most straightforward *comparative* analyses (such as those implicit within **Supplementary Table S1**) fraught with potential bias (Richardson and Speigelhalter 2020) – not least as a result of differences in the populations or contexts compared (Lourenço et al. 2020), and differences in the phase of any outbreak(s) therein (Sun et al. 2020).

For so-called *'predictive'* analyses – which rely on the individual and joint information available from optimally parameterised variables (covariates or 'predictors') to accurately estimate/classify a disease characteristic or variable of interest – the availability of sufficient data from carefully selected covariates (each offering tangible contributions to the model) is key (Arnold et al. 2020). In these models, sampling variation (and associated selection bias; Ellenberg 1994) is much less of a concern, except in as much as this might affect the external validity of the model's estimates/classifications (Fuller 2019). However, a critical weakness of these models is that they cannot be used for robust *causal inference* (i.e. to assess which of the included covariates

[3] Such 'predictions' include the estimation (or classification) of unknown, unmeasured or poorly measured/specified variables *either* retrospectively (or, at best, in near real time) *or* prospectively (in the future) based on the information available from other known/measured covariates (so-called 'predictors'). Both use statistical models (or 'algorithms') that have been 'trained' on datasets in which the 'predicted' variables have been (accurately) measured/specified. While the former better reflects 'interpolative estimation/classification' than 'prediction' in the literal sense, the latter generates 'literal predictions/extrapolations' that are nonetheless very different to the 'predictive *projections*' generated through model*ling* of the underlying *processes* theorised (or known) to be involved. In these, robust *causal* knowledge (both theoretical and empirical) is critical to the accuracy and precision their projections achieve. Widespread misunderstanding of the distinctions between these three forms of 'prediction' ('interpolative estimation/classification', 'literal prediction/extrapolation' and 'predictive projection'; see Figure S1) underpin their misapplication and misinterpretation, and fuel much of the bias – and many of the errors – that pervade contemporary epidemiology and may yet undermine the application of machine learning and AI therein (Arnold et al. 2020).[4,5]

[4] This is why epidemiological best practice should not rely on 'spotting' errors and biases, and should instead *assume* such problems are possible (if not likely), and diligently search for, root out and address these in the same way that parametricians routinely evaluate whether their data are normally distributed and homoscedastic (and thereby comply with two key assumptions of many parametric statistical models). Indeed, contemporary best practice extends the optimisation of parameterisation further by evaluating whether categorisation, transformation or interaction terms are required to maximise the (individual and joint) information that covariates provide to 'predictive' models; with the resulting models then subjected to repeated testing and evaluation. Similar diligence is required when selecting which variables to include (and which to exclude) from the 'covariate adjustment sets' required to minimise the risk of confounding while avoiding 'inferential mediator adjustment bias' in models that support robust causal inference. All such models benefit from careful parameterisation, as well as from a fuller understanding of the questions they can address (and those they cannot).

might act as genuine or important *causes* of, and candidates for intervention on, the characteristic/parameter predicted; Greenland 1996). Nevertheless, they are commonly used as if they *can* and thereby invoke a form of inferential bias known as the "Table 2 fallacy"[5] (Westreich and Greenland 2013) which can lead to fundamental misinterpretation of the causal relationships involved, with potentially dangerous implications for the focus and design of subsequent interventions (Eberhardt and Scheines 2007; Rehkopf et al. 2016).

Indeed, robust *causal inference* requires very different statistical models since, unlike those designed to optimise *'prediction'*, analyses capable of supporting causal inference can be extremely vulnerable to 'selection-collider bias' (or 'collider stratification bias'; Cole et al. 2010), which can invalidate the causal interpretation of statistical relationships observed within discrete subsets of any population (Munafò et al. 2018). They are also sensitive to inadequate/under-adjustment for potential confounders (VanderWeele and Arah 2007) and inappropriate/over-adjustment for mediators ('inferential mediator adjustment bias' or simply 'over-adjustment bias'; Schisterman et al. 2009; Richiardi et al. 2013) – both of which can strengthen, attenuate or even reverse the association observed between a speculative cause (or 'exposure') and its potential consequence (or 'outcome'; Cook and Ranstam 2017).[6]

Given such analytical and inferential biases remain commonplace amongst contemporary analyses of observational (i.e. non-experimental) clinical data (von Elm and Egger 2004; Blair et al. 2007; Detweiler et al. 2016; Pocock et al. 2004), it is not very surprising that they have also undermined so many preliminary analyses of COVID-19 data (see: Nussbaumer-Streit et al. 2020; Wynants et al. 2020). Indeed, this may be inevitable given that, at the start of the pandemic, when resources for diagnostic testing were in short supply, it made good sense to target those contexts where the identification of infected (and infectious) individuals was likely to provide the most benefit (Cheng et al. 2020; Pettit et al. 2020). But this then meant that the most accurate information has only been available for very specific subgroups within the population (such as health and social care practitioners, symptomatic individuals, and those with severe disease); and analyses on such subgroups run a significant risk of selection-collider bias – a bias that can be very difficult to detect, and can remain hidden, without subsequent analyses of more inclusive samples (Munafò et al. 2018).

While the 'Table 2 fallacy' and 'inferential mediator adjustment bias' are arguably easier to spot, both are more or less routine within clinical studies involving multivariable statistical analysis (hence Cook and Ramstam 2017).[7] They also affect a substantial proportion of more nuanced epidemiological analyses (Davey Smith and Ebrahim 2001); and whilst it is not the intention of this Commentary to highlight examples of poor practice, it is worth pointing out that these issues have undermined some of the most important (and potentially valuable/influential) studies undertaken during the COVID-19 pandemic thus far. These include: the OpenSAFELY Collaborative's study examining sociodemographic, behavioural and clinical correlates of COVID-19 mortality using health service data for almost 17.5 *million* adults in England (which nonetheless fell foul of the 'Table 2 fallacy'; Williamson et al. 2020); and a transatlantic consortium combining international sequencing data and local health service records to compare the severity of disease amongst patients infected

[5] This results from (mis)interpreting the coefficients of individual covariates within outputs from a single (one step) multivariable model (which are commonly those reported in a second Table, hence the fallacy's name) as evidence of their (independent) causal relationship with the predicted variable of interest. Instead these coefficients represent only the residual contribution each covariate makes to the model *after* adjustment for *both* the individual *and* joint information available from all other included/adjusted covariates – a residual contribution that can deviate in both size and direction from any true causal effect.

[6] 'Inferential mediator adjustment bias', which results from the inappropriate adjustment for mediators (variables falling on the causal pathway between the speculative cause/exposure and its potential consequence/outcome) in analyses intended to support causal inference, is the bias responsible for the 'Table 2 fallacy' (albeit, under those circumstances where the model in 'Table 2' was designed for *'prediction'* and subsequently repurposed/misinterpreted as a suitable basis for *causal inference*).

[7] Indeed, *none* of the empirical clinical studies (and only a handful of the epidemiological analyses) examined when preparing this Commentary appeared to recognise the important distinction between *'prediction'* and *causal inference* (and the different methodological considerations that each require).

with two emerging strains of SARS-CoV-2 (Korber et al. 2020; which nonetheless involved 'inferential mediator adjustment bias').[8]

Without access to the data involved in all such cases, it is not possible to establish the extent to which the biased coefficients (mis)represent the *true* causal relationships involved – or, if you prefer, whether George Box would classify these biases as 'mice' or 'tigers'. In some instances the likely consequences of some biases might be too trivial to require pressing attention (such as the implication that smoking is modestly protective against COVID-19 mortality – as the biased inference from Williamson et al.'s 2020 analyses suggest). But in others the severity of the risks involved (such as the possibility that mediator adjustment masked a genuine relationship between viral strain and disease severity in Korber et al.'s 2020 analyses) warrants urgent re-evaluation.

Simulation, extrapolation and projection: compartmental and agent-based modelling

Balancing the risks of error and bias against the conceptual and empirical insights available from *model organisms* and *descriptive statistics*, and from *analytical comparisons*, *'predictions'* and *causal inference*, becomes all the more challenging when these insights inform/underpin the *simulations*, *extrapolations* and *projections* used to forecast the future course of the COVID-19 pandemic. Indeed, the curves and waves these techniques generate (e.g. Kissler et al. 2020; Simeone 2020) have become both emblematic of, and instrumental in, the way we conceptualise disease progression; our expectations of what lies ahead; and the assumptions that both entail (Jones and Helmreich 2020). Yet such modelling takes a variety of different forms (Jewell et al. 2020) – including those based entirely on theory; those determined solely (and strictly) by the data; and those that fall someway in between – and it is worth examining each of these in detail to establish what questions they can (and can't) address (Holmdahl and Buckee 2020; Ioannidis et al. 2020; Siegenfeld and Bar-Yam 2020).

In practice, no modelling can be entirely theory-based since any such theory (and any parameterisation its application requires) relies at least in part on knowledge derived from real-world (empirical) studies of disease biology, transmission and progression (e.g. **Supplementary Table S1**). When such knowledge is shaky (as it has been during the early stages of the COVID-19 pandemic; Raleigh 2020; Weinberger et al. 2020) such modelling can generate projections that deviate markedly from the observed (or expected) numbers of infections, cases and deaths (Groen et al. 2020; Lourenço et al. 2020). And when the underlying theory is itself flawed, the projections may continue to diverge from reality even when the observed data to which these are compared have been collected and summarised competently, consistently and accurately (Luo 2020). Nonetheless, (predominantly) theory-based models do have considerable potential utility for *simulating* the likely impact of variation or changes in disease transmission and severity; and for *evaluating* hypothetical alternative interventions (such as those invoked to 'flatten' the curve during the COVID-19 pandemic; Adam 2020; Jewell et al. 2020; see **Supplementary Figure S2a**).

At the other extreme, exclusively data-driven modelling – which involves fitting curvilinear functions that best represent any apparent trend amongst the available data (IMHE and Murray 2020; Richardson and Speigelhalter 2020; Yang et al. 2020) – is also rarely used in the strictest ('data-driven') sense; even though recent advances in computational power and the availability of 'Big data' have made data-driven modelling techniques very popular in *'predictive analytics'*. As described earlier,[3] these make use of the 'joint information' provided in any given ('training') dataset to generate algorithms capable of estimating/classifying unknown, unmeasured or poorly measured/specified parameters/variables (Zhang et al. 2020) – algorithms that can then be refined through application in further ('testing') datasets (Li et al. 2020; Mondal et al. 2020). Although these techniques can struggle to generate accurate 'far-literal predictions' (by extrapolating the likely value of parameters/variables '*beyond*, and *far from*', the available data; Luo 2020), they can work well for 'near-literal predictions' (involving extrapolation '*beyond*, but *nearby*', the available data; Yadav 2020; Jewell et al. 2020); and work better still for 'predictive interpolation/classification' (and thereby improve the

[8] The authors of both these studies were well aware of these biases in advance of publication, and it is not clear why these biases were neither acknowledged nor competently addressed in the final versions they subsequently published.

estimation, specification or classification of parameters/variables *within* the available data; Mondal et al. 2020; Ellison et al. 2020; see **Supplementary Figure S2b**).

However, these advances in analytical technique are challenging to apply in data-driven forecasting of disease outbreaks, where the volume of available data (particularly at the outset) can be small and therefore vulnerable to chance variation, error and a small number of extreme values (Crozier 2020); and where – particularly for newly emerging diseases, like COVID-19 – the joint information from which such algorithms are generated only accumulates gradually over time, so that there is often limited scope for algorithmic training and testing on multiple, large (and complete) datasets (Adam 2020; Fuller 2020a; Sperrin et al. 2020). Indeed, at the beginning of a disease outbreak, where the numbers of cases and deaths are small in comparison to the at-risk population (of susceptible individuals), even a modest amount of variation, or a small degree of error, can have dramatic effects on key disease parameters and extrapolation based thereon (Crozier 2020; see **Supplementary Figure S2b**). These issues become less troublesome as the outbreak progresses and data accumulate, because curve-fitting functions can then reduce the influence of random variation and error in a similar fashion to that achieved using traditional 'smoothing' techniques (such as the 3-day rolling averages favoured by Richardson and Speigelhalter 2020). Nonetheless, any modelling that relies on measured values of key parameters/variables will be vulnerable to any changes in 'ascertainment' (the sampling, definition and identification of cases and deaths; Weinberger et al. 2020) which commonly accompany improvements in the accuracy, capacity and coverage of measurement – just as we have seen during the COVID-19 pandemic with successive improvements in the accuracy of diagnostic tests, and the extension of testing beyond high-risk groups and symptomatic individuals once testing capacity increased (Cheng et al. 2020; Pettit et al. 2020).

While changes in ascertainment can affect *all* modelling procedures, data-driven modelling is particularly susceptible to its effects because this approach relies entirely on the data to characterise the underlying 'data generating mechanism' on which insight into disease transmission and progression then relies (Tennant et al. 2017; Henley et al. 2020). When data are sampled and measured consistently over time, curve-fitting can be used to identify any signals offering evidence of disease transmission and progression dynamics that might be specific to the pathogen and the host population involved, or to the context in which pathogen-host interactions occur (Jewell et al. 2020). However, when changes in data sampling and/or measurement coincide with changes to the pathogen, host or host-pathogen relationship (such as the changes in exposure and transmission elicited through travel restrictions, social distancing, contact tracing and case-isolation), data-driven modelling can find it impossible to differentiate between these (e.g. IMHE and Murray 2020; Raleigh 2020). Under such circumstances, even data-driven modelling using AI's advanced curve-fitting algorithms offers little more than descriptive visualisations of the measurements available (Jones and Helmreich 2020), with scant insight into critical causal processes or likely future outcomes. Such modelling therefore retains only a residual degree of utility for estimating missing (or uncertainly/inexpertly ascertained) measurements through 'predictive interpolation/classification' (Ellison et al., 2020).

In practice, the vast majority of disease modelling techniques that epidemiologists and mathematical biologists use fall somewhere between those that are (strictly) theory-based and those that are (strictly) data-driven – what we might call 'theory-plus-data'-based modelling (which include 'compartmental' models as well as 'point process modelling' and 'agent-based modelling'; Bertozzi et al. 2020; Richardson and Speigelhalter 2020). Perhaps the best known of these involve theoretical/conceptual frameworks based on ostensibly plausible pathways of transmission between susceptible, (exposed,) infected and removed/recovered individuals (or discrete subsections of the population) – the so-called 'SIR' and 'SEIR' models which date back to the 1920s (e.g. Kermack and McKendrick 1927). These frameworks then use the best available empirical estimates of key disease parameters (such as the basic reproduction number, R_0; the serial interval; and the infection-fatality rate, IFR; SAGE 2020) to generate projections that can either: be simply compared to the empirical data (thereby identifying insightful discrepancies between the expected and observed numbers of cases and deaths); or fitted to match the empirical data by altering (or 'tuning'; Anastassopoulou et al. 2020; Simeone 2020; see **Supplementary Figure S2c**) one or more of the modelling parameters (thereby identifying, as before, insightful discrepancies between the best available estimates of

these parameters and those that best match the data). Such models share many of the modest advantages, but also many of the substantial disadvantages, of (mostly) theory- and (strictly) data-based modelling (Bertozzi et al. 2020). However, they arguably have greater utility for: highlighting critical discrepancies between the best available *estimates* of key disease parameters and the values of these *required* for the model to fit the data; and estimating critical parameters that vary over time and place, or are challenging to measure directly (such as the effective reproduction number, R_e; though see also: Petermann and Wyler 2020).

Conclusion

Clearly, epidemiologists have deployed an extensive range of models and modelling techniques to rapidly characterise many of the key parameters relevant to the transmission, severity and potential mitigation of COVID-19 (**Supplementary Figure S1**). Each of these techniques are based upon very different sets of assumptions; each are best suited to answer very different questions; and each require substantial thought (and technical expertise) to avoid analytical and inferential bias (Nussbaumer-Streit et al. 2020; Wynants et al. 2020). While all of these models are bound to be 'wrong', some will be 'useful' (Box 1976); and, together, the best of them offer complementary insights into the nature of the disease, and substantial foresight into what the past *might* mean (Probert et al. 2018) and the future *might* hold (Kissler et al. 2020; see **Supplementary Figures S2a-c**).

At present, the ability of models involving simulation, extrapolation and projection to deliver accurate projections (or 'far-literal predictions') of future events is severely constrained by: the (in)stability of the contexts in which they are applied; and how well the models actually reflect the underlying mechanisms and processes involved. In future, insight from analytical models best suited to strengthen estimation/classification (through 'predictive interpolation') and functional understanding (through 'causal inference') can play a pivotal role in: better characterising such contexts; and strengthening modelling of the underlying 'data generating mechanisms' on which accurate and reproducible projections (and perhaps even 'far-literal *predictions*') might then rely.[9]

In the meantime (and into the future), the epistemological value of all such models – for learning about, and understanding the nature of, any novel disease – will continue to rely on the thoughtfulness, diligence and competencies of those involved in their application, interpretation and dissemination (Sperrin et al. 2020). Substantial weaknesses in all three areas pose a significant and enduring challenge for epidemiology; and for the 'theory-free' application of new data-centric techniques (such as the algorithms developed using unsupervised machine learning and AI; Desai et al. 2019) where the potential for bias, misinterpretation, misunderstanding and misrepresentation remains but may be even harder to root out.

References

Adam A. 2020. The simulations driving the world's response to COVID-19. Nature. 580:316-8.

Anastassopoulou C, Russo L, Tsakris A, Siettos C. 2020. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PLoS ONE. 15: e0230405. DOI: 10.1371/journal.pone.0230405

Arnold KF, Davies V, de Kamps M, Tennant PW, Mbotwa J, Gilthorpe MS. 2020. Reflections on modern methods: generalized linear models for prognosis and intervention—theory, practice and implications for machine learning. Int J Epidemiol. Feb 26. DOI: 10.1093/ije/dyaa049

Ashofteh A, Bravo JM. 2020. A study on the quality of novel coronavirus (COVID-19) official datasets. Stat J IAOS. 36: 291–301.

Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Kalhori SR. 2020. Predicting COVID-19 incidence through analysis of google trends data in iran: data mining and deep learning pilot study. JMIR Public Health and Surveillance. JMIR Public Health Surveill 6:e18828. DOI: 10.2196/18828

[9] This view, that interpolative estimation/classification and causal inference are critical to accurate prediction (beyond mere extrapolation) is not without its detractors (Broadbent 2015), but will be persuasive to those adopting a more pluralistic positivist approach that charts a path between excessive scepticism and over-reliance on definitive evidence (Fuller 2020a).

Beggs L, Briscoe R, Griffiths C, Ellison GTH, Gilthorpe MS. 2020. Intervention differential effects and regression to the mean in studies where sample selection is based on the initial value of the outcome variable: an evaluation of methods illustrated in weight-management studies. Biostat Epidemiol. Mar 22. DOI: DOI: 10.1080/24709360.2020.1719690

Bertozzi AL, Franco E, Mohler G, Short MB, Sledge D. 2020. The challenges of modeling and forecasting the spread of COVID-19. Proc Nat Acad Sci. 117:16732-8.

Blair A, Stewart P, Lubin JH, Forastiere F. 2007. Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. Am J Indust Med. 50:199-207.

Blower SM, Dowlatabadi H. 1994. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. Int Stat Rev. 62:229-43.

Box GEP. 1976. Science and statistics. J Am Stat Ass. 71:791-9.

Broadbent A. 2015. Causation and prediction in epidemiology: A guide to the "Methodological Revolution". Stud Hist Philosoph Biol Biomed Sci. 54:72-80.

Burn-Murdoch J, Giles C. 2020. UK suffers second-highest death rate from coronavirus. Financial Times. May 28. https://www.ft.com/content/6b4c784e-c259-4ca4-9a82-648ffde71bf0 Accessed 2 Aug 2020.

Callaway E, Cyranoski D, Mallapaty S, Stoye E, Tollefson J. 2020. The coronavirus pandemic in five powerful charts. Nature. 579:482-483.

Cheng MP, Papenburg J, Desjardins M, Kanjilal S, Quach C, Libman M, Dittrich S, Yansouni CP et al. 2020. Diagnostic testing for Severe Acute Respiratory Syndrome–related Coronavirus 2: a narrative review. Ann Intern Med. Jun 2. DOI: 10.7326/M20-1301

Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, Poole C. 2010. Illustrating bias due to conditioning on a collider. Int J Epidemiol. 39:417-20.

Cook JA, Ranstam J. Statistical models and confounding adjustment. 2017. Br J Surg. 2017 Apr 104: 786–7.

Crozier C. 2020. Forecasting s-curves is hard. Apr 20. https://constancecrozier.com/2020/04/16/forecasting-s-curves-is-hard/ and https://vimeo.com/408599958 Accessed 2 Aug 2020.

Davey Smith G, Ebrahim S. 2001. Epidemiology—is it time to call it a day? Int J Epidemiol. 30:1-11.

Desai AN, Kraemer MU, Bhatia S, Cori A, Nouvellet P, Herringer M, Cohn EL, Carrion M, Brownstein JS, Madoff LC, Lassmann B 2019. Real-time epidemic forecasting: challenges and opportunities. Health Secur. 17:268-75.

Dorigatti I, Okell L, Cori A, Imai N, Baguelin M, Bhatia S, Boonyasiri A, Cucunubá Z, Cuomo-Dannenburg G, FitzJohn R, Fu H, Gaythorpe K, Hamlet A, Hinsley W, Hong N, Kwun M, Laydon D, Nedjati-Gilani G, Riley S, van Elsland S, Volz E, Wang H, Wang R, Walters C, Xi X, Donnelly C, Ghani A, Ferguson N. 2002. Report 4: severity of 2019-novel coronavirus (nCoV). Imperial College London, London. Feb 10. https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-02-10-COVID19-Report-4.pdf Accessed 2 Aug 2020.

Detweiler BN, Kollmorgen LE, Umberham BA, Hedin RJ, Vassar BM. 2016. Risk of bias and methodological appraisal practices in systematic reviews published in anaesthetic journals: a meta-epidemiological study. Anaesth. 71:955-68.

Ellenberg JH. Selection bias in observational and experimental studies. 1994. Stat Med. 13:557-67.

Ellison GTH. 2020. Introducing causal inference to the medical curriculum using temporal logic to draw directed acyclic graphs. J Univ Teach Learn Pract. 17: Forthcoming. medRxiv. Aug 4. DOI: 10.1101/2020.08.02.20166900

Ellison GTH, Wylde L, Deacon C, Wilson S, Shepherd A. 2020. 'Prediction' in the era of Big Data: opportunities and challenges for intelligence professionals. Cognitio. 94: 6-7.

Eberhardt F, Scheines R. 2007. Interventions and causal inference. Philosoph Sci. 74:981-95.

Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunubá Z, Cuomo-Dannenburg G, Dighe A, Dorigatti I, Fu H, Gaythorpe K, Green W, Hamlet A, Hinsley W, Okell LC, van Elsland S, Thompson H, Verity R, Volz E, Wang H, Wang Y, Walker PGT, Walters C, Winskill P, Whittaker C, Donnelly CA, Riley S, Ghani AC. 2002. Report 9: Impact of non-pharmaceutical

interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College London, London. Mar 16. https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf Accessed 2 Aug 2020.

Fuller J. 2019. The myth and fallacy of simple extrapolation in medicine. Synthese. May 23. DOI: 10.1007/s11229-019-02255-0

Fuller J. 2020a. Models v. evidence. Boston Review. May 5. http://bostonreview.net/science-nature/jonathan-fuller-models-v-evidence Accessed 2 Aug 2020.

Fuller J. 2020b. Why coronavirus death rates can't be summed up in one simple number. Conversation. Apr 10. https://theconversation.com/why-coronavirus-death-rates-cant-be-summed-up-in-one-simple-number-135758 Accessed 2 Aug 2020.

Funk S, Gilad E, Watkins C, Jansen VA. 2009. The spread of awareness and its impact on epidemic outbreaks. Proc Nat Acad Sci. 106:6872-7.

Goldstein JR, Atherwood S. 2020. Improved measurement of racial/ethnic disparities in COVID-19 mortality in the United States. medRxiv. Jun 23. DOI: 10.1101/2020.05.21.20109116

Greenland S. 1996. Basic methods for sensitivity analysis of biases. Int J Epidemiol. 25:1107–16.

Groen D, Zeka A, Hashmi IMQ. 2020. Six unknown factors in coronavirus models and how they could affect predictions. Conversation. Apr 2. https://theconversation.com/six-unknown-factors-in-coronavirus-models-and-how-they-could-affect-predictions-134855 Accessed 2 Aug 2020.

Henley SS, Golden RM, Kashner TM. Statistical modeling methods: challenges and strategies. 2020. Biostat Epidemiol. 4:105-39.

Holmdahl I, Buckee C. 2020. Wrong but useful—what covid-19 epidemiologic models can and cannot tell us. N Engl J Med. May 15. DOI: 10.1056/NEJMp2016822

IHME (IHME COVID-19 health service utilization forecasting team), Murray CJL. 2020. Forecasting the impact of the first wave of the COVID-19 pandemic on hospital demand and deaths for the USA and European Economic Area countries. medRxiv. Apr 26. DOI: 10.1101/2020.04.21.20074732

Ioannidis JP. 2020. Coronavirus disease 2019: the harms of exaggerated information and non-evidence-based measures. Eur J Clin Invest. 50:e13222. DOI: 10.1111/eci.13222

Ioannidis JPA, Cripps S, Tanner MA. 2020. Forecasting for COVID-19 has failed. International Institute of Forecasters. Jun 11. https://forecasters.org/blog/2020/06/14/forecasting-for-covid-19-has-failed/ Accessed 2 Aug 2020.

Jackson M. 2002. Disease and diversity in history. Soc Hist Med. 15:323-40.

Jewell NP, Lewnard JA, Jewell BL. 2020. Predictive mathematical models of the COVID-19 pandemic: Underlying principles and value of projections. J Am Med Ass. 323:1893-4.

Jones DS, Helmreich S. 2020. The shape of epidemics. Boston Review. Jun 26. https://bostonreview.net/science-nature/david-s-jones-stefan-helmreich-shape-epidemics Accessed 2 Aug 2020.

Khamsi R. 2020. Coronavirus is bad. Comparing it to the flu is worse. Wired. Feb 8. https://www.wired.com/story/coronavirus-is-bad-comparing-it-to-the-flu-is-worse/ Accessed 2 Aug 2020.

Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, Küchenhoff H, Tooze JA, Wallace MP, Kipnis V, Freedman LS. 2020. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1—Basic theory and simple methods of adjustment. Stat Med. 39:2197–2231.

Kermack WO, McKendrick AG. 1927. A contribution to the mathematical theory of epidemics. Proc Roy Soc A. Aug 1: 700-21. DOI: 10.1098/rspa.1927.0118

Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. 2020. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science. 368, 860–868.

Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N, Giorgi EE, Bhattacharya T, Foley B, Hastie KM, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI, McDanal C, Perez

LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC, Saphire EO, Montefiori DC, on behalf of the Sheffield COVID-19 Genomics Group. Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell. Jun 26 DOI: 10.1016/j.cell.2020.06.043

Lash TL, Fox MP, Fink AK. 2009. Applying Quantitative Bias Analysis to Epidemiologic Data. Springer-Verlag, New York. ISBN: 978-0-387-87960-4

Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. 2014. Good practices for quantitative bias analysis. Int J Epidemiol. 43:1969-85.

Lescure FX, Bouadma L, Nguyen D, Parisey M, Wicky PH, Behillil S, Gaymard A, Bouscambert-Duchamp M, Donati F, Le Hingrat Q, Enouf V. 2020. Clinical and virological data of the first cases of COVID-19 in Europe: a case series. Lancet Infect Dis. Mar 27. DOI: 10.1016/S1473-3099(20)30200-0

Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J. 2020. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology. 19 Mar. DOI: 10.1148/radiol.2020200905

Lourenço J, Paton R, Ghafari M, Kraemer M, Thompson C, Simmonds P, Klenerman P, Gupta S. 2020. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. medRxiv. Mar 26. DOI: 10.1101/2020.03.24.20042291

Lloyd-Smith JO, Funk S, McLean AR, Riley S, Wood JL. Nine challenges in modelling the emergence of novel pathogens. Epidemics. 2015. 10:35-9.

Luo J. 2020. Predictive monitoring of COVID-19. Data-Driven Innovation Lab. May 11 https://www.newsbeast.gr/files/1/2020/05/COVID19PredictionPaper.pdf Accessed 2 Aug 2020.

Mondal MRH, Bharati S, Podder P, Priya Podder P. 2020. Data analytics for novel coronavirus disease. Informatics Med Unlock. 20:100374. DOI: /10.1016/j.imu.2020.100374

Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. 2018. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 47:226-35.

Nussbaumer-Streit B, Mayr V, Dobrescu AI, Chapman A, Persad E, Klerings I, Wagner G, Siebert U, Christof C, Zachariah C, Gartlehner G. 2020. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. Cochrane Systematic Review. Apr 8. DOI: 10.1002/14651858.CD013574

Park M, Thwaites RS, Openshaw PJ. COVID-19: lessons from SARS and MERS. Eur J Immunol. 2020 50:308-16.

Paules CI, Marston HD, Fauci AS. 2020. Coronavirus infections—more than just the common cold. J Am Med Ass. 2020 323:707-8.

Petermann M, Wyler D. 2020. A pitfall in estimating the effective reproductive number R_t for COVID-19. medRxiv. May 12. DOI: 10.1101/2020.05.12.20099366

Petersen E, Koopmans M, Go U, Hamer DH, Petrosillo N, Castelli F, Storgaard M, Al Khalili S, Simonsen L. Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. 2020. Lancet Infect Dis. Jul 3. DOI: 10.1016/S1473-3099(20)30484-9

Pettit SD, Jerome KR, Rouquié D, Mari B, Barbry P, Kanda Y, Matsumoto M, Hester S, Wehmas L, Botten JW, Bruce EA. 2020. 'All In': a pragmatic framework for COVID-19 testing and action on a global scale. EMBO Mol Med. 12:e12634. DOI: 10.15252/emmm.202012634

Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, Kasten LE, McCormack VA. 2004. Issues in the reporting of epidemiological studies: a survey of recent practice. Br Med J. 329:883.

Probert WJ, Jewell CP, Werkman M, Fonnesbeck CJ, Goto Y, Runge MC, Sekiguchi S, Shea K, Keeling MJ, Ferrari MJ, Tildesley MJ. 2018. Real-time decision-making during emergency disease outbreaks. PLoS Comput Biol. 14:e1006202. DOI: 10.1371/journal.pcbi.1006202

Raleigh VS. 2020. Tackling UK's mortality problem: covid-19 and other causes. Br Med J. 369:m2295. DOI: 10.1136/bmj.m2295

Rehkopf DH, Glymour MM, Osypuk TL. 2016. The consistency assumption for causal inference in social epidemiology: when a rose is not a rose. Curr Epidemiol Rep. 3:63–71.

Rettner R. 2020. How Does the New Coronavirus Compare with the Flu? Scientific American. Feb 7. https://www.scientificamerican.com/article/how-does-the-new-coronavirus-compare-with-the-flu/ Accessed 2 Aug 2020.

Reyna VF, Brainerd CJ. 2008. Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. Learn Individ Diff. 18:89-107.

Richardson S, Spiegelhalter D. 2020. Coronavirus statistics: what can we trust and what should we ignore? Guardian. Apr 12. https://www.theguardian.com/world/2020/apr/12/coronavirus-statistics-what-can-we-trust-and-what-should-we-ignore Accessed 2 Aug 2020.

Richiardi L, Bellocco R, Zugna D. 2013. Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol. 42:1511-9.

Ritchie H, Ortiz-Ospina E, Roser M, Hasell J. 2020. COVID-19 deaths and cases: how do sources compare? Our World in Data. Mar 19. https://ourworldindata.org/covid-sources-comparison Accessed 2 Aug 2020.

Qin L, Sun Q, Wang Y, Wu KF, Chen M, Shia BC, Wu SY. 2020. Prediction of number of cases of 2019 novel coronavirus (COVID-19) using social media search index. Int J Environ Res Pub Health. 17:2365. DOI: 10.3390/ijerph17072365

SAGE (Scientific Advisory Group for Emergencies). 2020. Reasonable worst-case planning scenario. Mar 29. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/897509/S0089_Reasonable_Worst-Case_Planning_Scenario_-_29.03.pdf Accessed 2 Aug 2020.

Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, Keogh RH, Kipnis V, Tooze JA, Wallace MP, Küchenhoff H, Freedman LS. 2020. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 2—More complex methods of adjustment and advanced topics. Stat Med. 39:2232–63.

Schisterman EF, Cole SR, Platt RW. 2009. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiol. 20:488

Siegenfeld AF, Bar-Yam Y. 2020. What models can and cannot tell us about COVID-19 (pp. 1–3). New England Complex Systems Institute. Proc Nat Acad Sci. Jun 24. DOI: 10.1073/pnas.2011542117

Simeone R. 2020. Qualitative forecast and temporal evolution of the disease spreading using a simplified model and COVID-19 data for Italy. medRxiv. Jun 25. DOI: 10.1101/2020.06.22.20137133

Sperrin M, Grant SW, Peek N. 2020. Prediction models for diagnosis and prognosis in Covid-19. Br Med J. 369:m1464. DOI: 10.1136/bmj.m1464

Sun K, Chen J, Viboud C. 2020. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. Lancet Digit Health. 2020 Feb 20. DOI: 10.1016/S2589-7500(20)30026-1

Tennant PWG, Arnold K, Berrie L, Ellison GTH, Gilthorpe MS. 2017. Advanced modelling strategies: Challenges and pitfalls in robust causal inference with observational data. Leeds Institute for Data Analytics (LIDA), Leeds. ISBN: 978-1-5272-1208-4

Tennant PWG, Harrison WJ, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, Keeble C, Ranker LR, Textor J, Tomova GD, Gilthorpe MS, Ellison GTH. 2019. Use of directed acyclic graphs (DAGs) in applied health research: review and recommendations. Int J Epidemiol. Forthcoming. medRxiv. 2019 Dec 27. DOI: 10.1101/2019.12.20.19015511.

VanderWeele TJ, Arah OA. 2011. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. Epidemiology. 22:42–52.

von Elm E and Egger M. 2004. The scandal of poor epidemiological research. Br Med J. 329:868-9.

Wang CJ, Ng CY, Brook RH. 2020. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. J Am Med Ass. 323:1341-2.

Weinberger DM, Chen J, Cohen T, Crawford FW, Mostashari F, Olson D, Pitzer VE, Reich NG, Russi M, Simonsen L, Watkins A. 2020. Estimation of excess deaths associated with the COVID-19 pandemic in the United States, March to May 2020. J Am Med Ass Int Med. Jul 1. DOI: 10.1001/jamainternmed.2020.3391

Westreich D, Greenland S. 2013. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. Am J Epidemiol. 177:292-8.

Whiting K. 2020. Two experts explain what other viruses can teach us about COVID-19 – and what they can't. World Economic Forum. Mar 20.

https://www.weforum.org/agenda/2020/03/coronavirus-covid-19-mers-sars-experts/ Accessed 2 Aug 2020.

Williamson E, Walker AJ, Bhaskaran KJ, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby P, Cockburn J. 2020. OpenSAFELY: factors associated with COVID-19-related hospital death in 17 million patients. Nature. 8 Jul. DOI: 10.1038/s41586-020-2521-4

Woolf SH, Chapman DA, Sabo RT, Weinberger DM, Hill L. 2020. Excess deaths from COVID-19 and other causes, March-April 2020. J Am Med Ass. Jul 1. DOI: 10.1001/jama.2020.11787

Wynants L, Van Calster B, Bonten MM, Collins GS, Debray TP, De Vos M, Haller MC, Heinze G, Moons KG, Riley RD, Schuit E 2020. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. Br Med J. 369:m1328. DOI: 10.1136/bmj.m1328

Yadav RS. 2020. Data analysis of COVID-2019 epidemic using machine learning methods: a case study of India. Int J Inf Tech. May 19. DOI: 10.1007/s41870-020-00484-y

Yang S, Cao P, Du P, Wu Z, Zhuang Z, Yang L, Yu X, Zhou Q, Feng X, Wang X, Li W 2020. Early estimation of the case fatality rate of COVID-19 in mainland China: a data-driven analysis. Annals of translational medicine. Ann Transl Med. 8:128. 10.21037/atm.2020.02.66

Yee J, Unger L, Zadravecz F, Cariello P, Seibert A, Johnson MA, Fuller MJ. 2020. Novel coronavirus 2019 (COVID-19): Emergence and implications for emergency care. J Am Coll Emerg Phys Open. 1:63-9.

Zhang K, Liu X, Shen J, Li Z, Sang Y, Wu X, Zha Y, Liang W, Wang C, Wang K, Ye L. 2020. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography. Cell. 181:1423–33.

Acknowledgements

Mark Gilthorpe kindly provided extensive feedback on draft components of this Commentary, which also benefitted enormously from discussions with colleagues in the Leeds Causal Inference Research Group (particularly Kellyn Arnold, Laurie Berrie, Mark de Kamps, Wendy Harrison, John Mbotwa and Peter Tennant) and with Thea de Wet, Bob Mattes, Martin Shakespeare, Steph Leddington, Ian Scaysbrook, Andrew Shepherd, Richard Mason, Steve Braganza and David Moss.