The Strange Numbers of Covid-19

Annibale Biggeri and Andrea Saltelli***

** University of Firenze and University of Padova*
*** University of Bergen*

Abstract

Never as with the present pandemics, numbers and the attendant activities of measuring and modelling have taken centre-stage. Yet these numbers, often delivered by academicians and media alike with extraordinary precision, rely on a rich repertoire of assumptions, including forms of bias, that can significantly skew both the numbers per se and the trust we repose in them. We discuss the issue in relation to a particular case relative to the numbers on excess mortality during the first wave of the Covid-19 pandemic in Italy. We conclude with some considerations about the use of science at the science policy interface in situations where facts are uncertain, stakes high, values in dispute and decision urgent.

Keywords: Sensitivity auditing, Sensitivity analysis, Mathematical modelling, Epidemiology, Reproducibility, Post-normal science.

1. Introduction

Increasingly, we live immersed in numbers. Numbers possess their own reactivity (Espeland and Stevens 2008); they shape the real; they are performative, seductive (Engle Merry 2016), they generate paths for new numbers to be produced in a reinforcing feedback loop (Engle Merry 2016). At a deeper level, they "create the environment that justifies their assumptions" (O'Neil 2016: 29), and endow these who produce them with legitimacy (Porter 1995).

With COVID-19, this process has received a powerful acceleration, inter-alia throwing into the limelight one mode of production of numbers—mathematical modelling, till yesterday confined to the expert communities of developers and users. With COVID-19, modelling jargon, such as "flattening the curve" (Gross and Padilla 2020), has entered into public life. For some authors, the pandemic is operating a "domestication of modelling" (Montgomery and Engelmann 2020) whereby "COVID-19 is coming to be known in maths and models" (Rhodes et al. 2020: 253):

> With COVID-19, we see that maths and models have agency as drivers of social action, translating models into citizen science and advocacy. #FlattenTheCurve

Argumenta 7, 1 (2021): 97-107
ISSN 2465-2334

DOI 10.14275/2465-2334/202113.big

> entangles science into social practices, calculations into materialisations, abstracts into affects, and models into society (Rhodes et al. 2020: 256).

With new power, come new conflicts. “Wild-Ass Covid numbers”, cries Rush Limbaugh, who adds “The minute I hear anybody start talking about models and modeling, I blanch” (Pielke 2020). Models, already suspected to be the cloak used to hide normative visions (Romer 2015), have thus become even more politicized. If the eighteen century was the century we tamed probability (Hacking 1990) (but did we?), the twenty-first is could be remembered as the century we tamed mathematical modelling (but shall we?).

Did these model produce the right COVID-19 numbers? Here the opinions are divided. John P. Ioannidis and Nassim N. Taleb, while clashing on how to interpret the pandemic, agree on one thing: the failure in model-based forecasting (Pinson and Makridakis 2020). For Caduff (Caduff 2020), one needs to interpret the model based reaction to the pandemic keeping in mind politicians’ need to hide the systematic disinvestments in public health promoted by neoliberal policies - a famous report of the OECD (Organisation for Economic Co-operation and Development) warned in 2015 against excess hospital beds (OECD 2015). For Caduff, this reaction has fed into nervous media reporting, authoritarian longings and

> mathematical disease modelling—a flexible and highly adaptable tool for prediction, mixing calculations with speculations, often based on codes that are kept secret and assumptions that are difficult to scrutinize from the outside (Caduff 2020: 481).

Caduff (2020) points to the model of Imperial College (Ferguson et al. 2020) as responsible for having triggered a chain reaction of—in his view excessive—responses. We can add to Caduff that the model of the Imperial College was held responsible of “Jarring the U.S. and the U.K. to Action” (Landler and Castle 2020), while many complained about the non-easy accessibility/readability of the Imperial College model, made of thousands of lines of coding without much by way of comments. This model is agent-based, and hence intrinsically stochastic. Also for this, its reproducibility was questioned. According to the same Ferguson (2020), experts from GitHub and Microsoft supported his team check the code, which was then made available.

As per reproducibility, some researchers from Edinburgh University found a discrepancy of almost 80.000 deaths, even when using the same inputs and pseudo-random numbers for the Monte Carlo simulations. Apparently this was not due to the stochastic nature of the model, but to a bug in the code (see the discussion at https://bit.ly/2TknWR7).

Additionally, the journal Nature commented that the revised version of the program was approved by Codecheck, a project which awards “certificate of executable computation”(Eglen 2020; Singh Chawla 2020). Nature adds that other scientists have verified that the code is reproducible (Singh Chawla 2020). In the view of many, models such as the one used by Imperial College fall into the family of ‘Chameleon models’. A chameleon model:

> asserts that it has implications for policy, but when challenged about the reasonableness of its assumptions and its connection with the real world, it changes its

> colour and retreats to being a simply a theoretical (bookshelf) model that has diplomatic immunity when it comes to questioning its assumptions (Pfleiderer 2020: 85-86).

For a review of modelling issues in the age of COVID-19 see Saltelli et al. 2020b.

More reflexive approaches to model validation have been put forward (Eker et al. 2018; Funtowicz and Ravetz 1994; Ravetz 2003), which invite looking at a fuller set of attributes of modelling, seen as a process. This includes investigating the interests, the framings, and the expectations of the developers (Stirling 2010). A reflexive approach hunts for tacit assumptions, implicit biases in the use of a particular modelling approach, and instances of over-interpretation of the results. An example is offered by sensitivity auditing (Saltelli et al. 2013), and another by the use of model pedigrees as in NUSAP (Funtowicz and Ravetz, 1990a; van der Sluijs et al., 2005). These approaches see mathematical modelling as a tool among many to be used in the context of a participatory approach to the analysis of policies. Participation is achieved via what is called an 'extended peer community'. This formulation is due to Post-normal science (PNS), an approach to using science when facts are uncertain, decisions urgent, stakes high and values in dispute (Funtowicz and Ravetz 1994, 1993, 1990b).

In the present pandemic, the numbers of deaths and infections have taken centre-stage. For Emmanuel Didier (Didier 2020), when a set of numbers establishes itself, other possible numbers and stories are obscured. He worries that the erosion of civil liberties associated to the fight to the pandemic may thus come to be reckoned with too late, if at all.

For practitioners trained in the use of numbers, the precision of some model-based forecasts is staggering. Using cost-benefit analysis and the controversial concept of value of a statistical life, some authors conclude that social distancing in the US will lead to a net benefit of about $5.2 trillion (Thunstrom et al. 2020). While modellers are trained to understand—one would hope—the conditionality of model-based inference, model predictions are offered at face value (Saltelli et al. 2020a).

The complexities of the disease, and its rapid transformations, have also hit on another sore-point of the science-policy interface: the existing science reproducibility crisis (Ioannidis 2005; Saltelli 2018; Saltelli and Funtowicz 2017).

While the crisis is blamed by many, inter alia, on a publish or perish imperative (Edwards and Roy 2017), the pandemic accelerated by the need to produce results rapidly, and hence the dangers of non-reproducibility and the need to retract faulty papers. Studies on chloroquine and hydroxychloroquine published in the Lancet and NEJM were retracted by their own authors, when these were forced to admit that they had no access to privately owned data sources (Joseph 2020).

Scholars trained in the tradition of post normal science (Funtowicz and Ravetz 1993) wonder of this situation might eventually favour the emergence of a new model for the relation between science and society (Waltner-Toews et al. 2020). Such a model, or 'new covenant', would embrace PNS concept of an extended peer community, made of accredited experts as well as lay public, whistleblowers, investigative journalists, and the community at large. With the pandemic, the entire planet becomes on such community, "as the appropriate behaviour and attitudes of individuals and masses become crucial for a successful response to the virus" (Waltner-Toews et al. 2020). For (Waltner-Toews et al. 2020) this extended peer community is the opposite of a technocratic, number and model-based decision strategy.

Yet the present war on vaccines offers an interesting paradox. While the authorities still subscribe to a top down, broadcast model of science communication, the other side—that variously known as conspirationists or complotists, or vaccine-hesitants, appear to constitute precisely the extended kind of peer communities advocated for by PNS (Dotto et al. 2020). The democratization of expertise in the age of internet may happen to be blind to the quality of the expertise itself. For a discussion of the downside of such a democratization see also (Mirowski 2020).

Going back to the role of models, some warn against instances of modelling hubris. Here the dimension of a model is mistaken for its quality, and one observes a lack (or misapplication) of principles and practices for good model development (Saltelli et al. 2020a).

While numbers are produced ever more assertively by the responsible authorities and experts via a top down approach (Dotto et al. 2020), increased scepticism is manifested by the general public.

What went wrong? Is the public right to look at evidence-based policy with suspicion (Saltelli and Giampietro 2017), a fortiori when based—or including steps of—mathematical models (Saltelli and Funtowicz 2014)? Are the numbers of COVID-19 indeed 'strange'?

2. A Case Study

To help answering we present a brief review of the numbers on excess mortality during the first wave of the Covid-19 pandemic in Italy. Excess mortality is defined as the difference between the total number of deaths—all-causes mortality—and the expected number of deaths—i.e. the counterfactual number of deaths it would have been observed in absence of pandemic. It has been advocated as a better measure of impact because it is less affected by reporting bias. (Beaney et al. 2020) Indeed, the reporting of Covid-19 as cause of death depends on not uniquely defined coding rules, and this bias was more severe during the first wave of the pandemic. On international comparisons, this reporting bias was considered one of the reasons for the mortality gradient between countries—for example the higher counts for Italy (Pearce et al. 2020).

Excess mortality is an interesting indicator also because it permits to take into account indirect effects of the pandemic, a disease burden consequent to the overall societal response to the pandemic. Positive and negative effects could eventually sum up when considering the total number of deaths.

This appealing indicator however depends strongly on the calculation of the expected death counts, a modelling exercise with its inherent assumptions.

On June 2020, we retrieve 5 studies evaluating excess mortality during the first wave of Covid-19 in Italy (Table 1). On May 2020, the official statistics (National Institute of Health and Italian Statistical Institute) reported a Covid-19 death counts around 35,000ths.

Scortichini et al. presented an analysis of the excess mortality across the Italian provinces, stratified by sex, age group and period of the outbreak (Scortichini et al. 2020). The analysis was performed using a two-stage interrupted time-series design using daily mortality data for the period January 2015-May 2020. In the first stage, they performed province-level quasi-Poisson regression models, with smooth functions to define a baseline risk while accounting for trends and weather conditions and to flexibly estimate the variation in excess risk during the outbreak. Estimates were pooled among provinces in the second stage using a mixed-effects multivariate meta-analysis. In the period 15 February-15 May 2020, they estimated an excess of 47,490 [95% empirical confidence intervals (eCIs): 43,984 to 50,362] deaths in Italy.

	Period	Unit of analysis	Excess deaths (95% Confidence/Credibility Interval)	Method	Age stratification
Scortichini 2020	15/2- 15/5	Province	47,490 (43,984-50,362)	Time-series	YES
Blangiardo 2020	1/1- 31/4	Municipality	41,030 (35,600-42,099)	Space-time Bayesian	NO
Alicandro 2020	29/2- 31/5	All country	46,000 (uncertainty intervals not available)	Descriptive statistics	NO
Modi 2020	16/2- 9/5	Municipality	49,000- 53,000 (only estimates from two different modelling assumptions)	Synthetic Control	YES
Magnani 2020	1/3-15/4	Municipality	45,033 (42,761-45884)	Mortality rates	YES

Table 1. Italian studies on national Covid-19 excess mortality estimates (with uncertainty intervals) available in the literature in June 2020.

Blangiardo et al. (2020) used all-cause mortality weekly rates by municipality, based on the first four months of 2016-2019. They modelled municipality weekly trends for 2016-2019, separately for males and females, for each week and year using a Poisson distribution with age-adjusted expected number of cases as offset. They specified a Bayesian hierarchical model on the log mortality relative risk allowing a province-specific temporal trend. Finally, they included a smoothed effect on weekly temperature at municipality level. They then used the output from the model for 2016-2019 to predict the number of deaths expected for each week of the 2020 follow-up period. They estimated 41,030 excess deaths (95% credibility interval CrI 35,600;42,099).

Alicandro et al. (2020) used the number of deaths registered in the first six months of 2020 and compared it with that registered in the previous quinquennium. There was an over 50% excess total mortality in March and a 38% excess in April, corresponding to over 46,000 excess deaths in those two months.

Modi et al. (2020) performed a counterfactual time series analysis on 2020 mortality data from towns in Italy using data from the previous five years as control. Specifically, they constructed a counterfactual for every region, i.e. the expected mortality counts under the scenario that the pandemic had not occurred. They then compare this counterfactual with the reported total mortality numbers for 2020 to obtain an excess death rate. They estimated that the number of COVID-19 deaths in Italy is between 49,000 and 53,000 as of May 9 2020.

Magnani et al. (2020) analyzed data by region, sex and age, and compared to expected from 2015-2019. The reference daily mortality rates were computed dividing the 2015-2019 average by the corresponding population. Ninety-five percent confidence intervals (95% CI) were computed assuming a Poisson distribution of the observed deaths [21]. Five-day moving averages were used to reduce random variation in the graphical presentation. Within the municipalities studied, 77,339 deaths were observed in the period between March 1st to April 15th, 2020, in contrast to the 50,822.6 expected. The extrapolation to the total Italian population suggests an excess of 45,033 deaths in the study period.

Almost all studies used the five years 2015-2019—one study limits to 2016-2019—as comparison to estimate the counterfactual expected mortality. The methods varied ranging from times-series to Bayesian spatio-temporal, a few based on simple averaging among previous years death counts. What is surprising is the overall agreement, considering the lower-upper limits of the estimates the variation was 35,600-53,000. All estimates oscillated around the crude difference of 46,000 between the deaths counts on 2020 and the average death counts 2015-2019.

Do these calculations represent a clear cut evidence of excess mortality or they reflect a partially conscious adherence to a given narrative of the pandemic, confirming the reluctance of investigators to depart from previous results and existing narratives (Blastland et al. 2020)?

Blastland et al. (2020) argued that scientists were more prone to persuade than to inform. We would rather change to "scientists were willing to inform on what they were persuaded". To clarify this aspect we report a study performed by Biggeri et al. (2020) on excess mortality in Italy. They averaged by season and noticed that there was a reduction in mortality in Italy during January and February 2020 compared to the previous five years. Indeed, in winter 2019-2020 there was no influenza epidemic and therefore the population later exposed to the Covid-19 pandemic was more fragile than the average population of the previous five years. The amount of susceptible people was larger because it was not harvested during a milder winter. The naïve comparison with the same months of the previous years was then biased, because the populations were not comparable—the 2020 population being more frail. Biggeri et al. obtained, for each municipality, the posterior predictive distribution under a hierarchical null model. This allowed to take into account the natural variability of the phenomenon and avoid the rigid comparison of the observed number of deaths to a fixed number of expected counts. They calculated 25,700 (95%CrI 15,963; 51,045) excess deaths for the two months of March and April 2020. The position of the authors was to assume that small variations around the expected value of mortality should be considered natural and not be counted as excess mortality.

It is of interest to notice that the large part of results on excess mortality in the literature at that time were consistent with the upper limit of the credibility interval reported in Biggeri et al. The point is that there is an underreporting of the uncertainty in the statistical estimates of excess mortality and we provide an example from Italy. This attitude of the scientists reflects the partially conscious adoption of a pessimistic point of view. This is highlighted in our Italian example by the position of most estimates closed to the upper limit of the Biggeri study.

The difference between the point estimate of the Biggeri study and the others point estimates in the literature is what is called structural uncertainty or model uncertainty, which should be added to the sampling variability which is summarized by the confidence interval.

Moreover, if we consider a trade-off between false discovery and false non-discovery, most of the results in the literature correspond to a pessimistic figure about the pandemic. The false discoveries are the number of results falsely declared positive—i.e. the number of municipalities we declared to have experienced an excess mortality while this would have not been true—and the false non-discoveries are the number of results falsely declared negative—i.e. the number of municipalities we declared to have not experienced an excess mortality while this would have been true. Having decided a rule for rejection of the null hypothesis of no excess mortality, we automatically fix the number of false discoveries and non-discoveries. The trade-off depends on the fact that if we select a rule that limits the number of false discoveries we will pay a larger number of non-discoveries. The larger the figure of excess deaths the larger the number of false discoveries—i.e. we are concerned of not incurring in a larger rate of false non-discovery. A pessimistic position corresponds to a strong concern about a false non-discovery.

3. Discussion

Our Italian example shows that impact estimates varied due to different methodological choices.

It just seems that it is never ‘just the data’. As noted in a recent work “observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty” (Breznau et al. 2021). Political scientists Giandomenico Majone offered and interesting question—which could also work as a motivation for sensitivity analysis: “Are the results from a particular model more sensitive to changes in the model and the methods used to estimate its parameters, or to changes in the data?” (Majone 1989: 62).

In other words, the data ‘speak’, but so do the many hypotheses and choices made by the different investigators to obtain their confession.

A similar recent example concerned the Rt index. Even here a team of UK investigators found confidence interval which did not overlap (Scientific Advisory Group for Emergencies 2020). A comment by Spiegelhalter et al. (2021) was:

> It’s good to explore the same question through competing approaches. Many independent teams come up with different estimates of the reproduction number R, from which a committee has to come to a consensus. We return to George Box’s quote: “All models are wrong; the practical question is how wrong do they have to be to not be useful”. No model will be “correct”, and the quoted uncertainty interval […] should be taken with a pinch of salt, as it assumes the model is the truth.

4. Conclusions

The present work points to the opportunity for more reflexive approaches, both in the specific exercise of the art of modelling and in the broader topic of the use of evidence for policy making.

As per modelling work, it would appear that overall Covid-19 statistics seem to rely on a pessimistic assumptions (Blastland et al. 2020).[1] There is a tendency in research to reproduce previous results without exploring impact of underlying assumptions (Bailey 2017).

Where to look for a solution? For some observers, institutions, regulators and health authorities should attend to the suggestions made by PNS practitioners of new model for the relation between science and society (Waltner-Toews et al. 2020). As discussed, this approach should replace the top-down science communication model presently adopted by authorities, and in a sense parallel the strategies deployed by the skeptics themselves (Dotto et al. 2020). One can note that the present crisis of expertise in the age of internet parallels what happened because of the introduction of the movable-types press in the XV century. The printing machine brought to a drastic reduction of the power of the Church and to a long season of religious wars, ending in the second half of the XVII century. It is still early to say where the present crisis will lead. Seen in this context, the case of COVID-19 and of its conflicted numbers is perhaps a small episode, but it is instructive nevertheless.

References

Alicandro, G. et al. 2020, "COVID-19 Pandemic and Total Mortality in the First Six Months of 2020 in Italy", *Med Lav*, 111, 351-53, doi: 10.23749/mdl.v111i5.10786.

Bailey, D.C. 2017, "Not Normal: The Uncertainties of Scientific Measurements", *R. Soc. Open Sci.*, 4, 160600, doi: 10.1098/rsos.160600.

Beaney, T. et al. 2020, "Excess Mortality: The Gold Standard in Measuring the Impact of COVID-19 Worldwide?", *Journal of the Royal Society of Medicine*, 113, 329-34, doi: 10.1177/0141076820956802.

Biggeri, A. et al. 2020, "A Municipality-Level Analysis of Excess Mortality in Italy in the Period January-April 2020", *Epidemiol. Prev*., 44, 297-306, doi: 10.19191/EP20.5-6.S2.130.

Blangiardo, M. et al. 2020, "Estimating Weekly Excess Mortality at Sub-national Level in Italy during the COVID-19 Pandemic", *PLOS ONE*, 15, 10, e0240286, doi: 0.1371/journal.pone.0240286.

Blastland, M. et al. 2020, "Five Rules for Evidence Communication", *Nature*, 587, 362-64, doi: 10.1038/d41586-020-03189-1.

Breznau, N. et al. 2021, "Observing Many Researchers Using the Same Data and Hypothesis Reveals a Hidden Universe of Uncertainty", doi: 10.31222/osf.io/cd5j9.

[1] The study of Biggeri et al. 2020 does not provide the "most correct" estimate. There is not a most correct model (Saltelli et al. 2014). Pessimistic/Optimistic are relative to each other, the study of Biggeri et al. can be labelled as optimistic. We comment on the pessimistic assumption because most of the estimates tends to be higher comparing to those from Biggeri et al. 2020.

Caduff, C. 2020, "What Went Wrong: Corona and the World after the Full Stop", *Medical Anthropology Quarterly,* 34, 4, 467-87.

Didier, E. 2020, "Politique du nombre de morts", *AOC Anal. Opin. Crit.*, https://aoc.media/opinion/2020/04/15/politique-du-nombre-de-morts/.

Dotto, C. et al. 2020, "The "Broadcast" Model no longer Works in an Era of Disinformation" [WWW Document], *First Draft,* https://firstdraftnews.org:443/latest/the-broadcast-model-no-longer-works-in-an-era-of-disinformation/ (accessed 4.30.21).

Edwards, M.A. and Roy, S. 2017, "Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition", *Environ. Eng. Sci.*, 34, 51-61, doi: 10.1089/ees.2016.0223.

Eglen, S.J. 2020, "CODECHECK Certificate 2020-010", [WWW Document], *Zenodo*, doi: 10.5281/zenodo.386 5491.

Eker, S. et al. 2018, "Practice and Perspectives in the Validation of Resource Management Models", *Nat. Commun.* 9, 5359, doi: 10.1038/s41467-018-07811-9.

Engle Merry, S. 2016, *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking,* Chicago: University of Chicago Press.

Espeland, W.N. and Stevens, M.L. 2008, "A Sociology of Quantification", *European Journal of Sociol*ogy, 49, 401-36, doi: 10.1017/S0003975609000150.

Ferguson, N.M. 2020, "Microsoft and GitHub Working with the Imperial College", [WWW Document], *Twitter*, https://twitter.com/neil_ferguson/status/124183545 6947519492.

Ferguson, N.M. et al. 2020, "Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand", [WWW Document], *Imperial College London*, https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf.

Funtowicz, S. and Ravetz, J.R. 1994, "The Worth of a Songbird: Ecological Economics as a Post-normal Science", *Ecological Economics,* 10, 197-207, doi: 10.1016/0921-8009(94)90108-2.

Funtowicz, S. and Ravetz, J.R. 1993, "Science for the Post-Normal Age", *Futures*, 25, 739-55, doi: 10.1016/0016-3287(93)90022-L.

Funtowicz, S. and Ravetz, J.R. 1990a, *Uncertainty and Quality in Science for Policy*, Dordrecht: Kluwer.

Funtowicz, S. and Ravetz, J.R. 1990b, "Post-Normal Science: A New Science for New Times", *Scientific European*, 169, 20-22.

Gross, J. and Padilla, M. 2020, "Coronavirus Glossary: Flattening the Curve, Pandemic, Covid-19 and More ", *The New York Times*, March 18 2020, https://www.nytimes.com/2020/03/18/us/coronavirus-terms-glossary.html.

Hacking, I. 1990, *The Taming of Chance*, Cambridge: Cambridge University Press.

Ioannidis, J.P.A. 2005, "Why Most Published Research Findings Are False", *PLOS Medicine*, 2, 8, 696-701, doi: 10.1371/journal.pmed.0020124.

Joseph, A. 2020, "Lancet, NEJM Retract Covid-19 Studies that Sparked Backlash", *Statnews*, https://www.statnews.com/2020/06/04/lancet-retracts-major-covid-19-paper-that-raised-safety-concerns-about-malaria-drugs/.

Landler, M. and Castle, S. 2020, "Behind the Virus Report That Jarred the U.S. and the U.K. to Action", *The New York Times*, April 2.

Magnani, C. et al. 2020, "How Large Was the Mortality Increase Directly and Indirectly Caused by the COVID-19 Epidemic? An Analysis on All-Causes Mortality Data in Italy", *Int. J. Environ. Res. Public. Health*, 17, 3452, doi: 10.3390/ijerph17103452.

Majone, G. 1989, *Evidence, Argument, and Persuasion in the Policy Process*, New Haven: Yale University Press.

Mirowski, P. 2020, "Democracy, Expertise and the Post-Truth Era: An Inquiry into the Contemporary Politics of STS", *Academia.edu*, https://www.academia.edu/42682483/Democracy_Expertise_and_the_Post_Truth_Era_An_Inquiry_into_the_Contemporary_Politics_of_STS.

Modi, C. et al. 2020, "How Deadly is COVID-19? A Rigorous Analysis of Excess Mortality and Age-dependent Fatality Rates in Italy", *medRxiv*, doi: 10.1101/2020.04.15.20067074.

Montgomery, C. and Engelmann, L. 2020, "Epidemiological Publics? On the Domestication of Modelling in the Era of COVID-19", [WWW Document], http://somatosphere.net/2020/epidemiological-publics-on-the-domestication-of-modelling-in-the-era-of-covid-19.html/ (accessed 5.15.20).

OECD (Organisation for Economic Co-operation and Development) 2015, *Fiscal Sustainability of Health Systems Bridging Health and Finance Perspectives*, Paris: OECD Publishing, doi: 10.1787/9789264233386-en.

O'Neil, C. 2016, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Random House.

Pearce, N. et al. 2020, "Accurate Statistics on COVID-19 Are Essential for Policy Guidance and Decisions", *Am J Public Health*, 110, 7, 949-51.

Pfleiderer, P. 2020, "Chameleons: The Misuse of Theoretical Models in Finance and Economics", *Economica*, 87, 81-107, doi: 10.1111/ecca.12295.

Pielke, R. Jr. 2020, "The Mudfight Over 'Wild-Ass' Covid Numbers Is Pathological", *Wired*, https://www.wired.com/story/the-mudfight-over-wild-ass-covid-numbers-is-pathological/.

Pinson, P. and Makridakis, S. 2020, "COVID-19: Ioannidis vs. Taleb", [WWW Document], *Int. Inst. Forecas*t, https://forecasters.org/blog/2020/06/14/covid-19-ioannidis-vs-taleb/ (accessed 6.17.20).

Porter, T.M. 1995, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton: Princeton University Press.

Ravetz, J.R. 2003, "Models as Metaphors", in Kasemir, B., Jäger, J., Jaeger, C.G. and Gardner, M.T. (eds.), *Public Participation in Sustainability Science: A Handbook,* Cambridge: Cambridge University Press.

Rhodes, T. et al. 2020, "A Model Society: Maths, Models and Expertise in Viral Outbreaks", *Crit. Public Health*, 30, 253-56, doi: 10.1080/09581596.2020.1748310.

Romer, P. 2015, "Mathiness in the Theory of Economic Growth", *Am. Econ. Rev*, 105, 89-93, doi: 10.1257/aer.p20151066.

Saltelli, A. 2018, "Why Science's Crisis Should not Become a Political Battling Ground", *Futures*, 104, 85-90.

Saltelli, A. et al. 2013, "What Do I Make of your Latinorum? Sensitivity Auditing of Mathematical Modelling", I*nt. J. Foresight Innov. Policy*, 9, 213-34, doi: 10.1504/IJFIP.2013.058610.

Saltelli, A. et al. 2020a, "Five Ways to Ensure that Models Serve Society: A Manifesto", *Nature*, 582, 482-84, doi: 10.1038/d41586-020-01812-9.

Saltelli, A. et al. 2020b, "Five Ways to Make Models Serve Society: A Manifesto – Supplementary online material", *Nature*, 582.

Saltelli, A. and Funtowicz, S. 2017, "What Is Science's Crisis Really about?", *Futures*, 91, 5-11.

Saltelli, A. and Funtowicz, S. 2014, "When All Models Are Wrong", I*ssues Sci. Technol.,* 30, 79-85.

Saltelli, A. and Giampietro, M. 2017, "What Is Wrong with Evidence Based Policy, and How Can it Be Improved?", *Futures*, 91, 62-71, doi: 10.1016/j.futures.2016. 11.012.

Scientific Advisory Group for Emergencies 2020, "SPI-M-O: Consensus Statement on COVID-19", [WWW Document], GOV.UK., https://www.gov.uk/government/publications/spi-m-o-consensus-statement-on-covid-19-8-october-2020 (accessed 5.19.21).

Scortichini, M. et al. 2020, "Excess Mortality during the COVID-19 Outbreak in Italy: A Two-Stage Interrupted Time-Series Analysis", *Int. J. Epidemiol.*, 49, 1909-17, doi: 10.1093/ije/dyaa169.

Singh Chawla, D. 2020, "Critiqued Coronavirus Simulation Gets Thumbs up from Code-checking Efforts", *Nature*, 582, 323-24, doi: 10.1038/d41586-020-01685-y.

Spiegelhalter, D. and Masters, A. 2021, "Covid Vaccines Saved Lives in England, but Why Do Estimates Differ?", *The Guardian*, July 4, https://www.theguardian.com/theobserver/commentisfree/2021/jul/04/covid-vaccines-saved-lives-england-but-why-do-estimates-differ.

Stirling, A. 2010, "Keep it Complex", *Nature*, 468, 1029-31, doi: 10.1038/4681029a.

Thunstrom, L. et al. 2020, "The Benefits and Costs of Flattening the Curve for COVID-19", *SSRN Electron. J.,* doi: 10.2139/ssrn.3561934.

van der Sluijs, J.P. et al. 2005, "Combining Quantitative and Qualitative Measures of Uncertainty in Model-Based Environmental Assessment: The NUSAP System", *Risk Anal.*, 25, 481-92, doi: 10.1111/j.1539-6924.2005.00604.x.

Waltner-Toews, D. et al. 2020, "Post-Normal Pandemics: Why COVID-19 Requires a New Approach to Science", [WWW Document], *STEPS Cent. Blog*, https://steps-centre.org/blog/postnormal-pandemics-why-covid-19-requires-a-new-approach-to-science/