Standards for evidence in policy decision-making

Kai Ruggeri[1]*, Sander van der Linden[2], Claire Wang[3], Francesca Papa[4], Zeina Afif[5], Johann Riesch[6], James Green[7]

[1]Assistant Professor, Columbia University

[2]University Lecturer, University of Cambridge

[3]Vice President for Research, Evaluation, and Policy, New York Academy of Medicine

[4]Policy Analyst, Organisation for Economic Cooperation and Development

[5]Senior Social Scientist, World Bank

[6]Principal Research Scientist, Max-Planck-Institut für Plasmaphysik

[7]Chief Scientist, NASA

*Correspondence to: kai.ruggeri@columbia.edu.

Abstract: Benefits from applying scientific evidence to policy have long been recognized by experts on both ends of the science-policy interface. The COVID-19 pandemic declared in March 2020 urgently demands robust inputs for policymaking, whether biomedical, behavioral, epidemiological, or logistical. Unfortunately, this need arises at a time of growing misinformation and poorly vetted facts repeated by influential sources, meaning there has never been a more critical time to implement standards for evidence. In this piece, we present a framework to limit risks while also providing a reasonable pathway for applying breakthroughs in treatments and policy solutions, stemming the harm already impacting the well-being of populations around the world.

"For emphasis, I run some risk of overstatement." – Charles Lindblom, 1959

Introduction

There is growing demand for scientists to improve how they communicate evidence to decision-makers and the public (National Academies of Sciences, Engineering, and Medicine 2017). While finding common ground across scientific disciplines is often challenging (Johnson, 2013), effective science communication is crucial in assisting policymakers to design evidence-based interventions that will benefit entire populations. There is expanding investment into evidence-based practices. However, there remains substantial heterogeneity in standards for defining evidence across scientific fields and policy domains, which is especially a burden during crises such as the COVID-19 pandemic, where warnings have long been raised but were not fully heeded (Cheng et al., 2007). In this paper, we propose standard guidelines to support communicating evidence to policymakers. Such standards benefit scientific progress and policymakers while encouraging wider appreciation for empirical evidence.[1]

Evidence in policy

As of early 2017, all 50 US states and the District of Columbia demonstrate at least a modest level of integrating evidence into one or more policy domains (Pew-MacArthur, 2017). The absence of a common standard for identifying, defining, or integrating evidence into policy decisions, however, has resulted in substantial variability in how advanced these processes are.

With the Foundations for Evidence-Based Policymaking Act of 2018 now law in the United States, establishing such standards has immediate value. The "Evidence Act" involves a number of guidelines, notably influencing what policy areas are given priority, how information is disseminated, how agencies should aim to learn from evidence, and how to evaluate a range of policy actions. Yet, with the wide spectrum of content that can be treated as evidence, how best to identify reliable and appropriate sources will remain a challenge in government institutions. In spite of these challenges, increased emphasis on utilizing scientific insights presents a clear opportunity to improve standards for the application of evidence in policy.

Six decades ago, Lindblom (1959) outlined the opportunities and challenges of linking those insights from science to applications in policy, best characterized by the quote at the start of this manuscript. As outlined here, these challenges remain today, and due to COVID-19, have suddenly returned to the fore.

Formulating evidence-based policy for COVID-19

The COVID-19 pandemic poignantly illustrates the need for robust evidence-based policy. Some of the most critical questions for a generation are now pressed on leaders around the world. How should countries respond to effectively limit the spread of the coronavirus? Why have some interventions, in certain countries, had more success than others? What information can be trusted for implementing at scale?

Answering these questions is now particularly taxing, due to the conjunction of several factors unfolding on a global scale:

1. Over-supply of scientific evidence

[1] Given the urgency of the topic, here we present foundational arguments; a supplementary document with further references for each will be available at xxx.yyy/zzz.

2. Increasingly complex political processes
3. Rapid diffusion of information and misinformation (often through social media)
4. High level of uncertainties on many aspects, including reliability of available data and extent of cross-country comparability.

Behavioral science suggests that the policy interpretation of existing information can be particularly prone to biases in this context of scarcity of time and resources (Mullainathan and Shafir, 2013). Specifically for COVID-19, it is not just a gap in evidence of 'what works', but multiplying uncertainties for decision-makers due to not having sufficient time to find out. As a result, formulating evidence-informed policies appears to be most challenging right when we most need it, and countries are approaching the issue very differently.

Public policies to mitigate the spread of COVID-19 also have the potential to draw on behavioral insights, such as how to effectively encourage frequent hand washing, motivating individuals to distance themselves physically from others, ensuring widespread compliance with medical advice, and evaluating the mental health effects of long-term isolation. This requires a need for all forms of evidence to be classified systematically, separating what is viable from what is merely plausible, aesthetic, or novel (Smaldino & McElreath, 2016).

Without authoritarian intervention, South Korea drastically slowed the spread of the virus through unprecedented testing regimens, early physical isolation, and rapid tracing to quarantine the infected (Zastrow, 2020). Contrarily, the United Kingdom has witnessed widespread controversy and threats to hundreds of thousands of lives over its initial decision to delay actions, in part, based on fears of "behavioral fatigue" spreading throughout the population. This prompted a public letter signed by over 600 behavioral scientists in the UK to reconsider given the lack of sufficient evidence to support the concept (Chater, 2020). In Italy, the government initiated an *ad hoc* Technical Scientific Committee to refine lockdown measures on the basis of scientific recommendations (Protezione Civile, 2020).

If the British government had implemented a systematic framework such as the one described here, it would have become clearer that the evidence on "fatigues" (behavioral, media, isolation) is disparate at best, of mixed quality, and has a concerning lack of randomized controlled trials in support. Using our proposed THEARI rating system (introduced below), it would have been likely that experts would have considered the evidence between the stages of "empirical" and "applicable", yet far from "replicable" or "impactful". While innovation and new approaches to large-scale interventions will likely be necessary to combat the pandemic, the survival of entire populations in face of such crises should not rely on such limited information when better information is clearly available. This example illustrates how the lack of systematic assessment of evidence can impede optimal policymaking.

Of course, these concerns are not unique to the COVID-19 pandemic, so applications are possible on both immediate and broader fronts.

THEARI - A simple framework for standards of evidence in policymaking

To establish standards for evaluating evidence in policy contexts, we developed the Theoretical, Empirical, Applicable, and Replicable Impact rating system, (THEARI; Fig. 1). This five-tier system ranges from one (theory only) to five (impact validated) full stars. Its purpose is to provide guidance for scientists and policymakers to classify what qualifies as evidence and potential appropriateness for application.

THEARI rates a given insight by determining what evidence underlies it. Rather than requiring a policymaker to assess the evidence subjectively, or for researchers to champion their own work, the rating centers and standardizes the assessment of evidence. We recommend using the standard to inform decision-making by making ratings visible on journals as badges (Nosek et al., 2015) or retrospectively by external raters, such as those conducting a systematic review or policy briefs. It also aims to provide conceptual clarity in a context where heterogeneity in (or absence of) standards exists between locations, policy domains, and scientific disciplines. Where little information is available but a decision is necessary, it can be used to align related debates. Where an entire body of evidence including effective interventions is available, it can be used to identify the most robust insights available. We refer to evidence here as *scientifically produced insights or conclusions reported through peer-review or other recognized specialist dissemination channels,* though there are certainly other forms.

Consider the increasing use of social norms in behavioral policy as outlined in Figure 1. Initial papers defined a specific issue (suboptimal behaviors). Additional studies went further by identifying clear behavioral roots (observation of group behavior influences individual choice). Interventions were then proposed and tested, followed by replications. Further validation through successful trials across a number of domains and locations then facilitated systematic study of real-world impacts. It is not mandatory that each step be explicitly, discretely fulfilled to proceed to the next level; higher levels would, however, help assure that lower levels are met.

The amount and quality of evidence we have today on the effectiveness of social norms allows informed applications of such interventions to the COVID crisis. In particular, through decades of applications and replications (Cialdini, 2012), we have built sophisticated awareness of the related impact of descriptive social norms (what most people do) versus injunctive social norms (what most people think is the right thing to do). Evidence suggests that policymakers should not try to mobilize action against socially disapproved behavior by depicting it as frequent, as this might backfire by inadvertently installing a counterproductive descriptive norm in the minds of the public (Smerdon et al., 2019) - such as by sending the signal that hoarding toilet paper is *common* rather than *undesirable*.

Figure 1. The THEARI rating system.

Validation level	Rating	Description of standard for evidence	Example application: Social norms
Theoretical Argument or possible explanation stated		A scientifically-viable concept has been proposed but lacks empirical testing or validation. May come in the form of a descriptive theory, explanation of an issue, or a framework of a wider construct. Opinions may be treated as theory.	Published articles suggest that many social challenges may be the result of common behaviors that influence unwanted choices in a population. There is no direct test of this theory, nor any original data produced along with it.
Empirical Concept described but not utilized		Insights exist that identify and explain a given issue using valid measurement of observation or phenomenon. Eventually, it should include a move toward consensus on interpretations of robust study. May include non-successful interventions or lower-power studies, with increasingly converging conclusions as new data are generated.	Studies of similar methods on recycling, designated smoking areas, and text messaging while driving conclude that observing a negative behavior increases likelihood of eliciting the behavior, and vice-versa. Findings are largely correlational in nature.
Applicable Concept has been used to elicit effect		Effective intervention or application completed, in a controlled trial where possible. Measurement of processes and effects considered valid. Effect should demonstrate value for scientific insight and/or practice via reasonably-powered study. Ideally, the method was pre-registered for one or multiple studies.	A messaging intervention informs all members of a university told that reusable water bottles are the standard choice on campus. Sales of single-use bottles decrease on campus by 15%. Students bring reusable bottles to class an average of 12 times per semester.
Replicable Effect has been repeated independently		Valid and effective interventions produce converging conclusions through successful replication in terms of setting, procedure, and measurement. This is also a safeguard against errors (e.g., false positives) or bias tied to an individual study.	Multiple universities encourage reusable bottles by presenting this as the norm on campus. Students bring reusable bottles to class between 8 and 15 times on average per semester.
Impact Effect has been appropriately replicated in practice with measurable value in real world		Successful translation of insight applied at scale, producing consistent and validated effects in line with prior conclusions. Findings validated at the highest conceivable power (i.e., populations) through real-world testing and replication of effects in multiple settings. Standard approach to implementation, evaluation, and interpretation of data.	A city sends updated tax letters to all homes that have not paid, which claim that the majority of homes are already in compliance. The number of delinquent households decreases by 5% in the following 30 days. A similar method is associated with a moderate reduction in household energy use in a different city after a 90-day evaluation involving all addresses in community.

The system is meant to apply to visible ratings of a study for compilation of inputs in policy decisions. In practice, the rating would be applied to any published work as a header, footnote, or badge. Awarding five shaded stars is discouraged; the implication is that there should always be an opening for further research, even when – or perhaps especially when – validated impact has been achieved. Notably, there is no rating for opinions, commentaries, or editorials.

Common definitions of evidence are good for science

Standards for evidence are also important within scientific circles. In 1952, when Owen Storey proposed that the whistling noise heard in radio communications was due to plasma in the Earth's atmosphere, his argument was so heavily refuted that even his academic advisor suggested he drop the idea or risk being ridiculed as a scholar. In suggesting supersonic solar winds, Eugene Parker was similarly rebutted, and only when an eventual Nobel Laureate came to his defense was the initial manuscript published. Fortunately, as converging evidence validated these theories over time, they became cannon in science and practice. These unfortunate trajectories ultimately resulted in positive outcomes, but also created two major concerns. First, what valuable evidence has not been mobilized due to subjective treatment? Second, what inefficiencies have resulted from the same circumstances? To an extent, clearer standards for these would provide one possibility for improvement on both fronts.

Alternatively, consider current debate on the imminent threat of climate change: while substantial evidence has led to near-consensus in the scientific community, unsubstantiated denials of causes and impacts receive disproportionate attention (Cook et al., 2016). This imbalance harms scientific progress and stalls action addressing climate change (Lewandowsky et al., 2015). Similarly, failure to act on correct information in the context of COVID-19 will inevitably have implications for scientific progress, national security, and human survival.

Improving applications and replicability of evidence increases public trust in the discovery process of researchers (Nosek et al., 2015; Wingen et al., 2019). It also creates efficiency in policymaking processes by limiting reliance on arbitrary, competing opinions, which are unfortunately common in science and policy debates (Head, 2010*;* Howlett & Mukherjee, 2017). Standardizing evidence ratings of insights for policymaking helps counter false media balance and science denial by providing a common framework for using, rating, and referring to the weight of scientific evidence. Behavioral research finds that motivated reasoning is less likely to occur when people have to "give reasons" for why they support a particular position (Ballarini and Sloman, 2017), which rating systems such as THEARI facilitate.

Standards for ranking the progression of available information exist in many applied domains (irrespective of policy relevance). These are primarily for the purpose of drawing clear distinctions between what should and should not inform critical decision-making. For example, the Daubert standard for evidence in legal proceedings and Technology Readiness Levels (TRLs) in NASA, which established thresholds for when innovative tools are ready for widespread implementation. In the mid-1970s, TRLs were developed as a discipline-independent metric to allow more effective assessment of the maturity of new technologies, with detailed definitions first published in 1995. Abstraction within TRL allows a clear definition of the level of development relevant to many fields. Institutional adaptations now range from technology investment in the European Commission to the development of fusion reactor materials (Riesch et al., 2016). While TRL inherently emphasizes the highest rating should be expected, it encourages progress by demonstrating room for improvement when only lower values have been validated.

In medical contexts, standards are more common, as comparisons among multiple interventions or treatments are fundamental for decision-making. For example, GRADE (Grading of Recommendations Assessment, Development and Evaluation) is a framework to rate the quality of scientific evidence in systematic reviews (from very low to high) to help inform evidence-based

clinical guidelines. To evaluate the potential of clinical interventions, RCTs start out as high-quality and observational research as low quality, the GRADE approach then rates up or down based on the quality of the underlying evidence (e.g. risk of bias, effect-size, confounders, etc). However, systematic reviews themselves have a number of limitations, not least being that they cannot correct for errors in original studies, forcing an 'old dog, wrong trick' approach to policy choices (Ruggeri et al., 2016): aggregating poor quality data does not correct for poor quality. There is also recent evidence that collaborative replications may be more reliable for producing valid insights (Kvarnen et al., 2019). We take this directly into account with THEARI by highlighting that the best evidence requires multiple lines of investigation and a plurality of robust methods, not necessarily one over another.

Appealing to different methodologies, ranging from theoretical models to randomized controlled trials (RCTs), quasi-experiments, and laboratory research can enable public bodies to leverage the complementary strengths of these techniques (OECD, 2019). Policymakers rely on information from many sources to make decisions, which makes the communication of evidence as critical to them as it is to the general public (Doubleday & Wildson, 2012).

THEARI ratings provide a common language to assist all sides in understanding the level of evidence developed on a topic or from a single study, not to oversimplify critical nuance. By applying THEARI, opinions are not equated with empirical findings across scientific and policy domains and the framework ensures that a variety of perspectives is still considered alongside a consistent metric for evidence available. This frame aims to highlight that not all scientific contributions are created equal. While there is value in appealing to different lines of evidence, it is crucial to distinguish what specific additions each level of evidence will bring to policymaker toolkits, specifically at the five levels proposed.

Oversimplification, particularly given the types of evidence and other influences in policy may only erode trust between the public, researchers, and decision-makers (Head, 2010), as do failures in replication(Wingen et al., 2019). Each rating aims to bring structure to those discussions without overstating the weight of a single finding, instead providing a reference for categorizing relevant, available evidence.

Table 1. Strengths, weaknesses, and potential risks for applying standards for evidence.

Actionable strengths	**Practical limitations**	**Dangers to avoid**
Give standard for comparing evidence, regardless of current state	Does not specify a point where evidence is sufficient for a decision	Absolute thresholds that undermine open science or set unrealistic minimums, particularly where a decision is urgent or risks would become imminent
Accessible scale for expert and lay audiences	A simplified tool referring to likely complex topics cannot always result in policies backed by robust study from top academic journals	Using standards to mask misinformation or poorly designed studies
Anyone can reassess even if a score has been proposed	'Amount' of evidence may vary depending on context of application, such as urgency of need or disposition/bias of those evaluating - it is not always possible to have all the desired information	Static evaluations of evidence that do not acknowledge replication failures or adapt to new evidence
Systematic but practical ratings that can be updated over time	Is likely many effective interventions were trialed before substantial evidence was available on the issue, which creates ambiguity in rating	Ignoring conflicts of interest in funded studies, which may be presented in especially strong terms in support of a finding, thus objectively strong evidence if bias not considered
Unbiased by arbitrary thresholds	Does not consider participation or bias in policy or research, meaning inter-reliability is critical	Assuming scientific evidence is the only feature in policy decisions
No mandate for when to use, especially if decision is urgent	Difficult to compare between high impact, low evidence; low impact, high evidence	Purpose-driven research that lowers standards for discovery
Possible for retrospective, ex ante, and ex post assessment in application	Quality assessment of specific study rigor, especially analysis (Johnson, 2013), is separate but necessary and should emphasize quality, not volume	Interpreting a single rating as reflective of all features of a particular study – the rating should explicitly apply only to the primary insight

Going forward

Standards should be valuable to all members of the population, whether or not they value scientific evidence. Communicating those standards as well as the evidence is a major challenge (Broomell & Kane, 2017), especially in the context of health and medicine (Politi et al., 2007). In 2007, a team of researchers from Hong Kong (Cheng et al., 2007) published a warning letter about the re-emergence of SARS-like coronaviruses, and how it was a "time bomb" (p. 683). Their work, which they support with over 400 evidence-based references, would clearly meet the highest levels of evidence-based policy thresholds, and was backed by other studies of experts (Bruine de Bruin et al., 2006), yet the outbreak occurred with seemingly minimal preparation. This is not a specific fault of any single group, but using a simplified and informed standard for identifying the best quality evidence for policy action is again an urgent need.

In presenting THEARI, the ultimate benefit we envision is setting a common framework as a starting point for utilizing evidence in policy discussions, overcoming biases and the effects of inconsistent definitions or unreliable insights. This encourages policymakers to place more value on evidence by providing support for meaningful arguments that may otherwise be disregarded as incongruent with current thinking, even amongst scientists. Researchers can remain encouraged to continue study without overly emphasizing immediate application to the detriment of discovery, while also increasing understanding between those who may seek to utilize current insights. Doing so effectively should result in improved public policy approaches that ultimately serve the well-being of populations around the world.

For emphasis, we run some risk of oversimplification.

References

Ballarini, C., & Sloman, S. A. (2017). Reasons and the "Motivated Numeracy Effect.". In *Proceedings of the 39th annual meeting of the Cognitive Science Society* (pp. 1580-1585).

Broomell, S. B., & Kane, P. B. (2017). Public perception and communication of scientific uncertainty. *Journal of Experimental Psychology: General*, *146*(2), 286.

Bruine de Bruin, W., Fischhoff, B., Brilliant, L., & Caruso, D. (2006). Expert judgments of pandemic influenza risks. *Global Public Health*. 1(2), 179-194.

Cialdini, R. B. (2012). The focus theory of normative conduct. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), Handbook of Theories of Social Psychology. Vol.2 (pp. 295–312). London: Sage.

Cook, J., Oreskes, N., Doran, P. T., Anderegg, W. R., Verheggen, B., Maibach, E. W., ... & Nuccitelli, D. (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, *11*(4), 048002.

Chater, N., (2020, 16 March). People Won't Get 'Tired' Of Social Distancing – The Government Is Wrong To Suggest Otherwise. *The Guardian*. Available at: https://www.theguardian.com/commentisfree/2020/mar/16/social-distancing-coronavirus-stay-home-government (Accessed: 21 March 2020).

Cheng, V. C., Lau, S. K., Woo, P. C., & Yuen, K. Y. (2007). Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. *Clinical Microbiology Reviews*, *20*(4), 660-694.

Doubleday, R., Wilsdon, J. (2012). Science policy: Beyond the great and good. *Nature*, 485 (7398), 301-302.

Head, B. W. (2010). Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society, 29*(2), 77-94.

Howlett, M., Mukherjee, I. (2017). Policy design: From tools to patches. *Canadian Public Administration, 60*(1), 140-144.

Johnson, V. E. (2013). Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences*, *110*(48), 19313-19317.

Kvarven, A., Strømland, E., & Johannesson, M. (2019). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behaviour*, 1-12.

Lewandowsky, S., Oreskes, N., Risbey, J. S., Newell, B. R., & Smithson, M. (2015). Seepage: Climate change denial and its effect on the scientific community. *Global Environmental Change*, *33*, 1-13.

Lindblom, C. (1959). The science of muddling through. *Public Administration Review, 19*(2), 79-88.

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. Macmillan.

Munafò, M. R., & Davey-Smith, G. (2018). Robust research needs many lines of evidence. *Nature.* 553 (7689), 399–401.

National Academies of Sciences, Engineering, and Medicine. (2017). *Communicating Science Effectively: A Research Agenda.* The National Academies Press, Washington, DC.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Contestabile, M. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422-1425.

OECD. (2019). *Delivering Better Policies Through Behavioural Insights: New Approaches.* OECD Publishing, Paris.

Pew-MacArthur Foundation. (2017). *How states engage in evidence-based policymaking: A national assessment*. Pew Charitable Trusts, Washington, DC.

Politi, M. C., Han, P. K., & Col, N. F. (2007). Communicating the uncertainty of harms and benefits of medical interventions. *Medical Decision Making*, *27*(5), 681-695.

Protezione Civile, (2020). *Decree of the Head of Department n. 371, 5 February 2020, on the Institution of a Scientific Committee.* Accessible at: http://www.protezionecivile.gov.it/amministrazione-trasparente/provvedimenti/-/content-view/view/1206025

Riesch, J., Han, Y., Almanstötter, J., Coenen, J. W., Höschen, T. , Jasper,B., ... Neu, R. (2016). Development of tungsten fibre-reinforced tungsten composites towards their use in DEMO—potassium doped tungsten wire. *Physica Scripta*, (T167), 014006.

Ruggeri, K., Maguire, Á., & Cook, G. (2016). The "Next Big Thing" in Treatment for Relapsed or Refractory Multiple Myeloma May Be Held Back by Design—Between the Lines. *JAMA Oncology*, *2*(11), 1405-1406.

Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science*, *3*(9), 160384.

Smerdon, D., Offerman, T., & Gneezy, U. (2019). 'Everybody's doing it': on the persistence of bad social norms. *Experimental Economics*, 1-29.
Wingen, T., Berkessel, J. B., & Englich, B. (2019). No replication, no trust? How low replicability influences trust in psychology. *Social Psychological and Personality Science*, *11*(4), 454-463.

Zastrow, M. (2020, 18 March). South Korea is reporting intimate details of COVID-19 cases: has it helped? *Nature News.*

Prior version of THEARI from Ruggeri, K., Stuhlreyer, J., Immonen, J., Mareva, S., Paul, A., Robbiani, A., Thielen, F. Gelashvili, A., Cavassini, F., & Nairu, F. (2019). In K. Ruggeri (ed). *Behavioral insights for public policy: Concepts and cases.* Routledge.

FIGURE 3.2 Index for Evidence in Policy (INDEP).

Acknowledgments: We thank Bhaven Sampat (Columbia University), Thomas Zurbuchen (NASA), Marion Barthelemy (United Nations), Maja Friedemann (University College, London), Tomas Folke (Columbia University), and Tobias Wingen (University of Cologne) for input and coordination support on this manuscript.

Author contributions: KR was responsible for conceptualization, project administration, resources, writing (all drafts, all sections and features), editing, reviewing, and submission; SvdL was responsible for broad writing and reviewing; FP also provided broad writing and reviewing; CW provided extensive comments and edits; ZA reviewed and edited specific sections; JR was responsible for writing and reviewing specific sections and the THEARI descriptions; JG was responsible for responsible for reviewing drafts and supervising KR.

Author note: The following text was prepared originally on the importance of standards in evidence, building input iteratively over several years. It has been adapted to retain the broad scope while including multiple applications to the COVID-19 pandemic. This a blunt version of the topic; supplemental materials complement this with further references and commentary.

Funding: No funding was received in support of this work.

Competing interests: Authors declare no competing interests.

Data and materials availability: Not applicable.