Epidemics 47 (2024) 100775

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Scenario design for infectious disease projections: Integrating concepts from decision analysis and experimental design

Michael C. Runge [a,*,1], Katriona Shea [b,2], Emily Howerton [b,3], Katie Yan [b,4], Harry Hochheiser [c,5], Erik Rosenstrom [d,6], William J.M. Probert [e,7], Rebecca Borchering [b,8], Madhav V. Marathe [f,9], Bryan Lewis [f,10], Srinivasan Venkatramanan [f,11], Shaun Truelove [g,12], Justin Lessler [h,13], Cécile Viboud [i,14]

[a] U.S. Geological Survey, Eastern Ecological Science Center at the Patuxent Research Refuge, Laurel, MD, USA
[b] The Pennsylvania State University, University Park, PA, USA
[c] University of Pittsburgh, Pittsburgh, PA, USA
[d] North Carolina State University, Raleigh, NC, USA
[e] University of Oxford, Oxford, UK
[f] University of Virginia, Charlottesville, VA, USA
[g] Johns Hopkins University, Baltimore, MD, USA
[h] The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[i] Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

ARTICLE INFO

Keywords:
Scenario modeling
Design of experiments
Sensitivity analysis
Value of information
Multi-model projections

ABSTRACT

Across many fields, scenario modeling has become an important tool for exploring long-term projections and how they might depend on potential interventions and critical uncertainties, with relevance to both decision makers and scientists. In the past decade, and especially during the COVID-19 pandemic, the field of epidemiology has seen substantial growth in the use of scenario projections. Multiple scenarios are often projected at the same time, allowing important comparisons that can guide the choice of intervention, the prioritization of research topics, or public communication. The design of the scenarios is central to their ability to inform important questions. In this paper, we draw on the fields of decision analysis and statistical design of experiments to propose a framework for scenario design in epidemiology, with relevance also to other fields. We identify six different fundamental purposes for scenario designs (decision making, sensitivity analysis, situational awareness, horizon scanning, forecasting, and value of information) and discuss how those purposes guide the structure of scenarios. We discuss other aspects of the content and process of scenario design, broadly for all settings and specifically for multi-model ensemble projections. As an illustrative case study, we examine the first 17 rounds of scenarios from the U.S. COVID-19 Scenario Modeling Hub, then reflect on future advancements that could improve the design of scenarios in epidemiological settings.

* Corresponding author.
E-mail address: mrunge@usgs.gov (M.C. Runge).
[1] ORCiD: 0000–0002-8081–536X
[2] ORCiD: 0000–0002-7607–8248
[3] ORCiD: 0000–0002-0639–3728
[4] ORCiD: 0000–0002-8381–504X
[5] ORCiD: 0000–0001-8793–9982
[6] ORCiD: 0000–0002-7755–0559
[7] ORCiD: 0000–0002-3437–759X
[8] ORCiD: 0000–0003-4309–2913
[9] ORCiD: 0000–0003-1653–0658
[10] ORCiD: 0000–0003-0793–6082
[11] ORCiD: 0000–0002-0874–8692
[12] ORCiD: 0000–0003-0538–0607
[13] ORCiD: 0000–0002-9741–8109
[14] ORCiD: 0000–0003-3243–4711

https://doi.org/10.1016/j.epidem.2024.100775
Received 14 August 2023; Received in revised form 4 April 2024; Accepted 20 May 2024
Available online 24 May 2024
1755-4365/Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Epidemics prompt many questions, from public health policy makers wanting to know how to intervene, to members of the public wanting to know what to expect, to industrial leaders wanting to know how to react. These questions are all necessarily forward looking, creating a demand for epidemiologists to project what may happen in the future. Epidemiological models provide a way to integrate historical observations, biological and sociological knowledge, and our understanding of disease mechanisms to produce projections of epidemiological outcomes into the future. These projections can be used to guide decisions (by governments, industries, and individuals) about how to respond, and to guide research investment (to reduce uncertainty in projections or processes).

Quantitative scientists make a distinction between forecasts and scenario projections (Reich et al., 2022). Forecasts are unconditional predictions about the future, statements about what is expected to happen. The most useful forecasts are probabilistic, expressly recognizing and transparently quantifying the uncertainty in the prediction. Scenario projections, on the other hand, are conditional predictions about the future, statements about what would happen if a set of conditions were to be met. The most useful scenario projections are also probabilistic, but conditionally so; they typically express the probability of outcomes if certain conditions are met, but do not usually quantify the probability of those conditions themselves being met. Scenario projections, then, are exploratory—they allow the examination and contrast of multiple futures. Both forecasts and scenario projections make assumptions (for example about social behavior or public health policy) and so, in a sense, the predictions from both are conditional on those assumptions. But in forecasts, the intent is to embed any uncertainty in those assumptions in the forecast itself, so that the forecast is a singular probabilistic representation of what is likely to happen. In scenario projections, some of the conditions are of specific interest and are isolated as separate scenarios.

Scenario modeling is common in many disciplines, including climate science (Krey, 2014), conservation biology (Nicholson et al., 2019), wildlife and fisheries management (Johnson et al., 1997), economics (McDowall and Eames, 2006), transportation (Bartholomew, 2007), urban planning (Khakee, 1991), energy development (Leung and Yang, 2012), agriculture (Pfister et al., 2011), invasion ecology (Shea et al., 2005), military planning (Dowse, 2021), disaster planning and response (Tyszkiewicz et al., 2012), nuclear war and terrorism (NASEM, 2023), and many others. One of the most visible global examples is the Coupled Model Intercomparison Project (CMIP), which has produced six phases of climate projections based on shared scenario specifications (Eyring et al., 2016; Meehl et al., 2000). The most recent phase of projections (CMIP6) is based on a set of "shared socio-economic pathways" and provides central evidence for the sixth assessment report of the Intergovernmental Panel on Climate Change. Importantly, the shared socio-economic pathways (and the "representative concentration pathways" of CMIP5) represent forcing scenarios (notably concerning carbon emissions); the climate projections are conditional on the scenario assumptions, but the likelihood of those conditions occurring was not estimated. Similar types of scenario projections are made in many other fields, but there is not yet a common lexicon that unites the large literature on this subject. In an influential book, Martelli (2014) argued that the field of scenario planning faces a number of shortcomings, notably a lack of clarity in the conceptual foundations, methodological inconsistency, and absence of evidence of effectiveness. A more recent review finds progress toward a synthesis of concepts and methods and increasing evidence of effectiveness, but notes that the field remains fragmented (Cordova-Pozo and Rouwette, 2023).

The use of scenario modeling has become pervasive in infectious disease epidemiology over the last two decades. Notable examples include modeling of different types and layers of interventions to control emerging outbreaks such as foot-and-mouth disease in the United Kingdom (Tildesley et al., 2006), avian influenza (Longini et al., 2005), the Ebola outbreak in West Africa (Meltzer et al., 2014), and the COVID-19 pandemic (Borchering et al., 2023; Borchering et al., 2021; Hellewell et al., 2020; Truelove et al., 2022; Walker et al., 2020). Additional notable use cases include the roll-out of new interventions for endemic pathogens where, for instance, scenario projections can help anticipate the benefits and dynamic changes associated with new vaccines or improved drugs (Flasche et al., 2016; Pitzer et al., 2009; Eaton et al., 2012; Houben et al., 2016). In some cases, the scenario projections are produced from a single model (e.g., Meltzer et al., 2014), while in others, the scenario projections come from multiple models (Flasche et al., 2016; Houben et al., 2016), drawing on a growing literature documenting the value of multi-model efforts (Johansson et al., 2019; Shea et al., 2020, 2023; Cramer et al., 2022; Prasad et al., 2023). Scenario design plays an important role in infectious disease projections over long time scales, not only to contrast different intervention schemes, but also to control for uncertainty in key parameters that may be magnified over time. However, there is little guidance on how to optimize scenario assumptions to answer particular public health questions, especially in the context of multi-model efforts.

Across fields, one of the central features of scenario projections is that "scenarios seem to exist in sets and the scenarios that inhabit those sets are systematically prepared to co-exist as meaningfully different alternatives to one another" (Spaniol and Rowland, 2019). How, then, are these sets developed? Many methods for scenario design exist and attempts have been made to classify the methods into several schools of approach (Amer et al., 2013). In this paper, we draw from the fields of decision analysis and experimental design to propose a framework for scenario development that integrates the three schools discussed by Amer et al. (2013). We place this work in the context of epidemiological modeling, but intend the framework to be more broadly useful. Our primary thesis is that clarity about the purpose of the scenarios is central to their design, and we offer a taxonomy of design purposes.

2. Purposes of scenario design

We approach scenario design like experimental design. First, a scenario design, like an experimental design, should have a purpose—a question (or questions) that the designers seek to answer. Second, a scenario design consists of a set of alternative scenarios (analogous to experimental treatments), which differ with regard to one or a few factors. Third, the foundational experimental design concepts of control, randomization, and replication have analogs in scenario design. The scenarios can be designed to control certain factors by prescribing shared assumptions or parameter values. By randomly sampling from the probability distributions for uncontrolled parameters, inferences from scenario comparisons can be extended to the full parameter space represented by those distributions. Controlling for these otherwise unspecified variables is common in experimental design through a method called pairing; this is also possible in scenario modeling by pairing replicates across scenarios (i.e., compare replicates with the same uncontrolled parameters, that therefore differ only by scenario). Further, each scenario can be replicated many times, either by soliciting repeated projections from a single model structure or by soliciting projections from multiple models of varying structure. These concepts are embedded in our framework for scenario design.

One of the central questions in scenario design is how the individual scenarios differ. Many factors that will affect future dynamics are unknown at the time of projection (e.g., human behavior or key aspects of pathogen biology). Scenario design, then, is the process of strategically choosing among those many uncertainties to identify a set of scenarios that together can achieve the purpose of the scenario projection exercise. Inspired by multiple rounds of COVID-19 projections that have addressed public health goals at different stages of the pandemic, we identify three primary purposes in scenario design: making decisions, exploring uncertainty, and identifying how decisions may be affected by

uncertainty. To understand the differences between these purposes, we distinguish two types of factors (often described as "scenario axes"): interventions (or decision options) and uncertainties. Interventions are factors that are under the control (or partial control) of one or more decision makers, such as vaccination policies, non-pharmaceutical intervention (NPI) policies, or hospital staffing and capacity. Uncertainties are factors that are not under any decision maker's control, but that might affect the outcomes or possibly even the choice of intervention. This distinction between intervention and uncertainty factors is not always sharp, and can depend on the primary audience. For instance, the arrival of a new virus variant will always be considered uncertain, as its emergence is beyond anyone's control. However, other factors, such as vaccine coverage, are more complicated, as they can be affected by the informational campaigns of public health agencies (a decision) and the behavioral responses of individuals (an uncertainty). Further, a factor that is an intervention for one decision maker (e.g., vaccination recommendation by the Centers for Disease Control and Prevention, CDC) might be an uncertainty for another decision maker (e.g., a hospital complex). The explicit purpose of a scenario design and its intended primary audience, however, can help shed light on whether a factor should be treated as a decision or an uncertainty.

2.1. Purposes of scenario design: a taxonomy

Scenario modeling is an attempt to glimpse something about the future, often with the intention of informing actions in the present. In this sense, there is a decision-making element to scenario modeling, but the decisions can have many purposes: to change the trajectory of the future through interventions; to respond to future outcomes; or to seek more information. We believe that understanding the purpose of a scenario modeling exercise informs the design of the scenarios, as well as any subsequent ability to evaluate the success of the exercise. We propose six classes of scenario design that stem from the three primary goals. Two-factor designs are very common in scenario modeling across all fields, so we provide shorthand for each class based on what the two factors would be, but note later that simpler (one factor) or more complex designs are possible.

2.1.1. Decision making (decision×decision in a 2×2 matrix design)

In a decision-making setting, the scenarios are designed to contrast alternative interventions, actions that are intended to influence the outcomes being modeled. For example, in Fig. 1(A), the scenario design consists of three scenarios, varying spatial extents of an intervention (nowhere, in 1 of 3 candidate locations, everywhere). If multiple types of interventions are being considered, a factorial arrangement of the levels of each might be of interest, so the scenario design could be, say, a 2×2 matrix with both axes being interventions. The interventions could be alternatives being considered by a decision maker (e.g., Borchering et al., 2023) or potential interventions being explored to nudge decision makers to consider new options (Meltzer et al., 2014).

2.1.2. Sensitivity analysis (uncertainty×uncertainty in a 2×2 matrix design)

The purpose of sensitivity analysis is to understand the contributions of different sources of uncertainty to the outcomes of interest (Saltelli et al., 2004), and potentially whether they interact. In a sensitivity analysis setting, then, the scenario axes focus on uncertainties. For example, in the 2×2 scenario design in Fig. 1(B1), one axis captures uncertainty about the basic reproductive number (R_0 of 2.2 or 3.0) and the other captures uncertainty about the serial interval (3 or 4.5 days). Note that there is no decision explicit in this design, although research efforts could be devoted to the more influential factor.

2.1.3. Situational awareness (uncertainty×uncertainty)

Three additional classes of scenario design resemble sensitivity analysis designs, in that the scenario axes focus on uncertainties only,

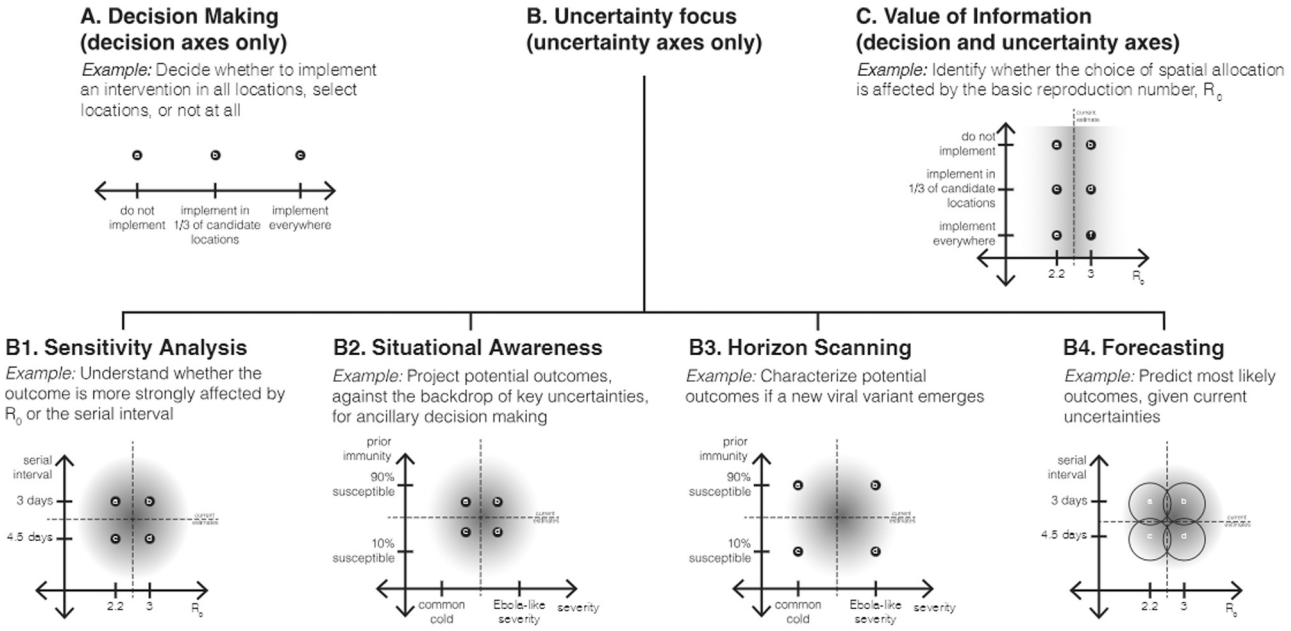

Fig. 1. Graphical depiction of six classes of scenario design, with heuristic examples. (A) In a Decision Making scenario design, the axis or axes are variables that are under the control of the decision maker; the purpose of the design is to understand the outcomes associated with different interventions. (B) Four classes of design have a similar structure (uncertainty axes only) but different purposes. (B1) Sensitivity Analysis designs focus on understanding the role of different sources of uncertainty on the outcomes of interest. (B2) Situational Awareness designs resemble Sensitivity Analysis designs, but have an additional purpose to provide insight about potential outcomes that may be relevant for ancillary decisions. (B3) Horizon Scanning designs explore the edges of the epistemic uncertainty, often to prompt insights about what could happen in the future, in an effort to develop new interventions. (B4) Forecasting designs postulate multiple hypotheses in the parameter space, with an appropriately weighted average of outcomes constituting a well-calibrated forecast, given the current uncertainty. (C) A Value of Information (VOI) design (decision axis×uncertainty axis) examines whether a source of uncertainty affects the relative effects of interventions. The dashed lines represent the current point estimates and the shaded regions represent the current confidence intervals for the uncertainty parameters.

but their purposes differ, possibly affecting construction of the scenarios. Sometimes scenarios are used for situational awareness, to give decision makers and the public a sense of the current state-of-the-world and what might be coming. In this way, different from the sensitivity analysis class, decisions are implied, although their effects are not embedded in the design. For example, in the case of an emerging disease, initial data may provide an estimate of the degree of prior immunity from cross-immunity and the severity of the disease, but there might be considerable uncertainty about both parameters (Fig. 1(B2)). A scenario design that examines plausible values of both parameters allows scientists to better understand the role of the sources of uncertainty, but also allows decision makers to anticipate the range of possible outcomes, and so consider how to prepare or respond.

2.1.4. Horizon scanning (uncertainty×uncertainty)

There is a large literature, primarily outside of epidemiology, that focuses on horizon scanning (also often called scenario planning; Sutherland and Woodroof, 2009). In this approach, scenarios are designed to explore plausible extremes of what could happen (Fig. 1 (B3)), as a way to provoke awareness of future possibilities, motivate preparation, avoid or plan for surprises, and encourage creation of new intervention strategies. Even though there are no interventions on the axes of this design, this approach is more decision-centric than sensitivity analysis or situational awareness, in that a decision maker is aware of looming threats and is looking for insight to guide novel interventions. In conservation settings, horizon scanning around the possible impacts of climate change has become an important approach, as natural resource management agencies realize that their old tools may no longer be effective in changing ecosystems. In epidemiological settings, the horizon scanning class has been used particularly in thinking about emergence of novel pathogens or variants, like spillover of avian influenza to humans (Colizza et al., 2007).

2.1.5. Forecasting (uncertainty×uncertainty)

The fourth approach with a focus on uncertainty aims to design a set of scenarios that can be combined into an unconditional probabilistic forecast of the future, by careful choice of scenarios to bracket key uncertainties (Fig. 1(B4)). This approach differs fundamentally from the other types of designs described above: first, the set of scenarios needs to collectively represent the full degree of uncertainty about influential parameters (e.g., those included in analytical expressions for the reproductive numbers derived from mechanistic models); and second, the likelihood of the individual scenarios needs to be specified (or derived from experience). With these conditions, a weighted combination of the scenario projections forms a proper forecast with appropriate uncertainty. The belief weights on the scenarios (i.e., likelihood of each scenario) can be updated dynamically in time as new evidence comes in, using a Bayesian approach or its generalization, Dempster-Shafer Theory (Shafer, 1990). In the field of natural resource management, when such dynamic scenario forecasting is embedded in a Markov decision process, it is called “adaptive management” (Chadès et al., 2012; Walters, 1986); similar approaches are commonly used in machine learning and artificial intelligence applications (Sutton and Barto, 2018). Thus, while descriptions in the epidemiological literature tend to make a sharp distinction between forecasts and scenario projections, we believe the latter can be transformed into the former with appropriate design and weighting.

2.1.6. Value of information (decision×uncertainty)

Value of information (VOI) is a common concept in the field of decision analysis that assesses whether the more effective intervention (rather than its outcome) is sensitive to the uncertainty. Thus, value of information is a form of sensitivity analysis from the standpoint of the decision maker (Felli and Hazen, 1998). In a value of information design, at least one scenario axis is an intervention and at least one other is an uncertainty. For example, in Fig. 1(C), the design consists of 6 scenarios, in a 3×2 design, with one decision axis (the same spatial implementations as in Fig. 1(A)) and one uncertainty axis (R_0); the result of particular interest would be whether the ranking of decision options was different under the two values of the basic reproductive number. From a decision-making perspective, this is the most important design, because it evaluates intervention alternatives while also investigating whether their performance is robust to major sources of uncertainty; it also can be used to inform the value of gathering more information prior to making a decision.

There are other uses of scenario modeling, in training and tabletop exercises, where the users’ interactions with the scenarios are central to their purpose. The use of scenario modeling has a rich history in military training (Straus et al., 2019; Kim et al., 2014; NRC, 2008), as well as other fields. The goals of tabletop exercises include understanding inter-agency coordination, preparedness in terms of personnel, equipment, and protocols, and other aspects of complex responses to emerging threats. Policy and decision makers are assigned roles and asked to make various decisions during an evolving scenario. Examples of settings where scenario modeling has been used in tabletop exercises include responses to a novel SARS-like agent (Dausey et al., 2005), release of plague bacteria (*Yersinia pestis*, Henderson et al., 2001), and a new outbreak of foot-and-mouth disease in the United Kingdom (DEFRA, 2018). The design of scenarios in these types of exercises tends to be more complex than the others described above, with nested and branching scenarios that respond to user actions. The details of such designs are beyond the scope of this paper, although many of the elements that we discuss will be relevant.

3. Scenario designs used by the COVID-19 Scenario Modeling Hub

To illustrate the proposed scenario classification, we retrospectively analyzed 17 rounds of scenario designs developed by the U.S. COVID-19 Scenario Modeling Hub (SMH) (https://covid19scenariomodelinghub.org/). Since December 2020, the SMH has convened multiple modeling teams to generate scenario-based projections of COVID-19 cases, hospitalizations, and deaths over 3–24 month horizons, in close collaboration with U.S. public health agencies. The 17 rounds of scenarios available for study addressed different needs at different stages of the pandemic. The scenario classification described in this paper (Fig. 1) was not available when SMH scenarios were designed, but we have applied it retrospectively, recapturing the intent of each round through publicly released reports and internal notes taken during the design process (Table 1). Scenarios were typically designed through an iterative discussion process between the SMH coordination group, participating modeling teams, and public health partners. This process took anywhere from 3 to 86 days (median 32.5 days; see also Loo et al., 2024 for more details). For each round, we have identified the “motivating audience” as the decision-making body that the designers had foremost in mind. In most rounds, the motivating audience was the collection of federal, state, and local public health agencies with authority to set public health policy or guidance, but in 6 of the 17 rounds, the design was more strongly motivated by consultation with a specific public health partner (notated in Table 1 with bold type). It is important to note, however, that the SMH coordination group always worked with, and had in mind, the needs of multiple decision-making agencies beyond the motivating audience, and, at times, these considerations also influenced the scenario design.

Of the 17 rounds, 3 were classified as Decision designs (decision×decision), 6 rounds as VOI designs (decision×uncertainty), and the other 8 as some form of uncertainty×uncertainty design (Table 1). The Decision designs were clustered earlier in the pandemic (December 2020-March 2021), representing a period when decisions regarding NPIs and vaccines were most needed. The VOI designs occurred throughout the pandemic (January 2021-April 2023) and focused on various vaccination decisions, such as increase of primary series

Table 1
Retrospective determination of scenario designs used by the U.S. COVID-19 Scenario Modeling Hub (SMH) in its first 17 rounds, released February 2021-April 2023. Most SMH rounds included 4 scenarios reflecting different levels of controls or epidemiological situation, depending on the stage of the epidemic. All rounds, except for Rounds 1 and 17, were organized as a 2×2 table representing two axes or key epidemic drivers (e.g., vaccination and NPI, Factors 1 and 2 columns), with a high and low value assumed for each of these drivers. The classification of scenario design (last column) arises from considering the axes types, as well as the purpose of the round: a "Decision" class arises when all scenario axes are decision axes; a value of information (VOI) class arises when a decision axis is crossed with an uncertainty axis; and the remaining types (sensitivity analysis, situational awareness, and horizon scanning) arise from uncertainty by uncertainty structures. The classification of the axis types was done with reference to the motivating audience, but other audiences could use the results for other purposes. Audience abbreviations: *PHA*, public health agencies, that is, federal, state, or local public health decision makers (*default generic audience* for most rounds); CDC, U.S. Centers for Disease Control and Prevention; WH, the White House COVID-19 Task Force; ACIP, the CDC Advisory Committee on Immunization Practices; SMH, the U.S. COVID-19 Scenario Modeling Hub (for internal insights). Other abbreviations: NPI, non-pharmaceutical intervention (social distancing, masking, etc.); vax, vaccination.

Round #	Data Cut-off Date	Purpose	Motivating Audience	Factor 1 (axis type)	Factor 2 (axis type)	Scenario Design
1	12/15/2020	Examine the impact of several combinations of vaccination and NPI levels	*PHA*	Vaccination and NPIs[a] (decisions)	(none)	Decision
2	1/23/2021	Examine the impact of vaccination and NPIs in light of the emergence of a new variant (Alpha)	*PHA*	Vaccination and NPIs[b] (decisions)	Variant (uncertainty)	VOI
3	3/06/2021	Compare the effects of vaccination and NPIs, in light of limitations in vaccine distribution	*PHA*	Vaccination (decision)	NPIs (decision)	Decision
4*	3/27/2021	Compare the effects of vaccination and NPIs, with updated estimations of variant (Alpha) characteristics and availability of vaccine	**CDC**	Vaccination (decision)	NPIs (decision)	Decision
5	5/01/2021	Understand potential long-term outcomes, in light of ongoing reduction in NPIs and slower-than-expected update of vaccination	*PHA*	NPI (uncertainty)	Vaccination (uncertainty)	Situational awareness
6	5/29/2021	Understand potential long-term outcomes, in light of emergence of a new variant (Delta) and slowing vaccine uptake	*PHA*	Vaccination/Hesitancy (uncertainty)	Variant (uncertainty)	Situational awareness
7*	7/03/2021	(Same as Round 6, with updated variant parameters)	*PHA*	Vaccination/Hesitancy (uncertainty)	Variant (uncertainty)	Situational awareness
8[X]	8/14/2021	Examine the potential effects of waning immunity	*PHA*	Protection level after waning immunity (uncertainty)	Speed of waning immunity for natural infection and vaccination (uncertainty)	Sensitivity analysis
9	9/11/2021	Estimate the impact of childhood vaccination, with a stress test involving emergence of a new variant	**ACIP**	Childhood vaccination (ages 5–11) (decision)	Hypothetical variant (uncertainty)	VOI
10[X]	11/13/2021	Examine booster coverage aspirations needed to offset waning immunity	*PHA*	Waning immunity (uncertainty)	Booster coverage (decision)	VOI
11	12/18/2021	Understand medium- to long-term outcomes, as related to uncertainties about the Omicron variant	*PHA*, SMH	Variant Severity (uncertainty)	Immune escape/transmissibility[c] (uncertainty)	Situational awareness
12*	1/08/2022	Evaluate the potential impact of the Omicron wave (with parameters updated from Round 11)	*PHA*	Severity (reduction relative to Delta) (uncertainty)	Immune escape (uncertainty)	Situational awareness
13	3/12/2022	Examine the long-term impact of waning immunity, in the face of potential emergence of a new variant	**CDC**	Waning immunity (uncertainty)	Hypothetical variant (uncertainty)	Horizon scanning
14	6/04/2022	Compare the effects of different age targets for fall 2022 boosters, with a stress test involving a new variant	**ACIP**	Age targets for boosters (decision)	Variant (uncertainty)	VOI
15	7/30/2022	Examine the effect of the timing of rollout of bivalent boosters, with and without emergence of a new high immune-escape variant	**WH, ACIP**	Timing of reformulated boosters (decision)	Variant (uncertainty)	VOI
16	10/29/2022	Understand the 6-month outcomes, as related to uncertainty about booster update and the epidemiology of emerging variant "swarms"	*PHA*	Booster uptake (uncertainty)	Variant (uncertainty)	Situational awareness
17	4/15/2023	Examine the effect of different booster targets on multi-year outcomes, in the face of uncertainty about antigenic drift	CDC, ACIP	Age targets for new boosters[d] (decision)	Rate of immune escape (uncertainty)	VOI

* Rounds marked with an asterisk were primarily an update of the previous round.

[a] Round 1 included 4 scenarios that looked at combinations of social distancing measures, masking guidance, vaccine availability, and vaccine hesitancy, and did not include a factorial structure.

[b] The first axis of Round 2 was composed of three decision variables: state orders with regard to NPIs (social distancing, masking) and vaccine availability. A full factorial for the first axis would have had 8 levels (2x2x2), but this was reduced to just two levels by combining the most optimistic and most pessimistic choices for each factor.

[c] Round 11 had an implicit 2x2x2 design (severity of Omicron infection, transmissibility, and immune escape), but the second two factors were condensed together (low transmissibility coupled with higher immune escape; high transmissibility coupled with low immune escape). Thus, a fractional factorial design was used to allow exploration of more factors with fewer scenarios.

[d] Round 17 was a 3×2 design, where the first axis included three levels.

[X] These rounds were not publicly released.

coverage among adults (Round 2), expansion of the vaccine program among children (Round 9), or comparison of different booster strategies (Rounds 14 and 15). In these VOI designs, the second axes typically described properties of virus variants, extent of waning, or immune escape. Of the 8 rounds classified as uncertainty×uncertainty designs, 6 were considered as situational awareness, 1 as sensitivity analysis, and 1 as horizon scanning. Situational awareness rounds were designed to anticipate the arrival of new variants, or evaluate the potential impact of

growing vaccine hesitancy and declining NPIs. The round classified as sensitivity analysis was devoted to understanding the impact of waning assumptions on disease dynamics (training Round 8 in summer 2021, which was not publicly released). The round classified as horizon scanning explored potential interactions between waning immunity and a hypothetical immune escape variant in the post-Omicron period. Although SMH scenarios span many of the designs presented in our proposed classification, forecasting scenarios were not represented per se, in part because the primary purpose of the SMH was not to explicitly combine scenarios (see Bay et al., 2024 for a post-hoc application of this concept).

These classifications were challenging to make because the SMH rounds were used (and implicitly designed for) many audiences, each of which might interpret a design differently. For example, Round 4 was designed specifically with the CDC in mind in Spring 2021, and, at the time, the degree to which they should emphasize vaccination versus compliance with NPIs was important (Decision category, decision×decision axes). A similar classification would apply from the lens of a state or county public health agency, because recommendations to the public about vaccination and NPIs were in their authority. In contrast, for a hospital administrator, these scenarios might have served as situational awareness (uncertainty×uncertainty) that was useful in anticipating staff and resource needs over the coming months. Relatedly, Rounds 11 (December 2021) and 12 (January 2022) addressed the Omicron variant and were designed primarily for situational awareness, with scenarios informed by early data on variant characteristics from South Africa. However, given the limited amount of information available on Omicron severity in Round 11, a broad range of severity assumptions was chosen, so that this round could also be considered horizon scanning.

Another interesting challenge in retrospectively classifying the designs of the 17 SMH rounds was judging whether a particular axis was a decision or uncertainty axis, as the same axis designed for the same user might have had a different meaning at different stages of the pandemic. For instance, in the first four rounds, we interpreted the vaccination and NPI axes as decision axes, because the CDC and other public health agencies were actively grappling with how aggressively to recommend vaccination, how to allocate initially limited doses of vaccine, and how strongly to implement and enforce NPIs. In Round 5, however, by May of 2021, public health agencies seemed to have become somewhat resigned to the behavioral choices of individuals regarding vaccination and compliance with NPIs, and so we treated those factors as uncertainties rather than decisions. On reflection, it would have been easier to classify the axes and the scenario designs in the moment, and in consultation with the motivating audiences.

4. Detailed considerations in scenario design

In the following sections, we describe the elements of a scenario design in an epidemiological setting, more elements of experimental design that are pertinent, and other practical considerations. To guide the reader, we have provided a figure that recapitulates the process of scenario design and highlights key components that need to be considered (Fig. 2). When relevant, we illustrate these considerations with examples drawn from the 17 rounds of SMH scenario designs.

4.1. Elements of a scenario design

Several key elements should be considered in scenario design, including output metrics, details of how the scenarios differ, initial conditions, common factors, and the approach for handling uncertainty not otherwise expressed across scenarios. All such design decisions should be informed by the scenario's purpose. Transparency and precision is important for communication, especially in multi-model settings to ensure that contributing models produce comparable outputs.

4.1.1. Output metrics

A scenario design needs to specify output metrics to be projected. This choice can strongly influence the insights gained as well as interventions recommended (Probert et al., 2016). For infectious disease projections, output metrics might include incident or cumulative cases, hospitalizations, or deaths, which may be further broken down by subgroups (e.g., age, region, race, or ethnicity). Metrics that are not strictly epidemiological may also be of interest (e.g., business closure duration for NPI-based scenarios, Shea et al., 2023). Generally, multiple outputs are assessed separately, but multi-criteria decision analysis (Keeney and Raiffa, 1993) would permit composite outputs to be addressed. The scenario design also needs to specify the time frame and spatial extent of outputs, including temporal and spatial resolution. Another important consideration is the establishment of shared "ground truth" data for output metrics, particularly in the context of a collaborative effort, as these data are typically used for model calibration.

Details about how to summarize and report results also should be specified; this will depend in part on the desired inference. A full probability distribution for the output metrics might be desired, or summaries like the mean, median, or an exceedance probability might suffice. Alternatively, individual replicates (e.g., daily or weekly simulations) may be useful in several cases. First, for outputs that are time-dependent, presenting an average over replicates can mask important features, like the size of a peak in cases (Juul et al., 2021); calculation of the desired metric by replicate before averaging can help avoid this issue. Second, individual replicates can be used to contrast projections across scenarios (see discussion about pairing replicates in section 4.2 *Principles of experimental design as applied to scenarios*). Finally, individual replicates also offer more flexibility than summary measures, as the need for nuanced metrics may arise after scenarios have been designed. We refer the reader to Sherratt et al. (2024) for deeper comparison of summary outputs and individual replicates.

4.1.2. Distinguishing features of scenarios

The key aspects that differentiate the scenarios are the choice of axes and levels set for each (Fig. 2). As we discussed earlier, the choice of scenario axes arises out of the purposes of the scenario design, but even with clarity of purpose, the choice can be difficult. Notably, the sources of uncertainty may be quite numerous, and selecting the focal uncertainties can thus be challenging. A large factorial design (over many different axes of uncertainty) is possible to imagine but may be time-consuming to implement and even harder to interpret and communicate. Selection of a subset of uncertainties essentially asks the designers to conduct an implicit sensitivity analysis to identify the sources they think will most affect the outcomes of interest. This sensitivity analysis can be, and sometimes is, undertaken with a simplified prototype model; in other cases, the expert judgment of the designers is used to select the axes. A similar concern can arise with decision axes—there can be many different interventions and combinations of interventions to consider. In our experience, however, the decision axes are easier to identify, because they often arise from direct questions being posed by decision makers.

Levels on a decision axis may represent specific discrete interventions (e.g., approve a vaccine for a specific age group or not) or represent a continuous variable that is closely tied to a decision (e.g., high and low coverage of a vaccination campaign). In the latter case, the levels chosen might bracket the range of effects that could be achieved under different strategies. Sometimes counterfactual scenarios are used to evaluate the population-level benefits of an intervention. For example, at the beginning of COVID-19 vaccine rollout in December 2020, SMH Round 1 considered a scenario without any vaccination as well as scenarios with various coverage levels (Table 1). Because the vaccine had already been approved and manufacturing was in progress (FDA, 2020), the no-vaccination scenario was not expected to eventuate, but it was important for comparative purposes.

Levels on an uncertainty axis may be based on available estimates of

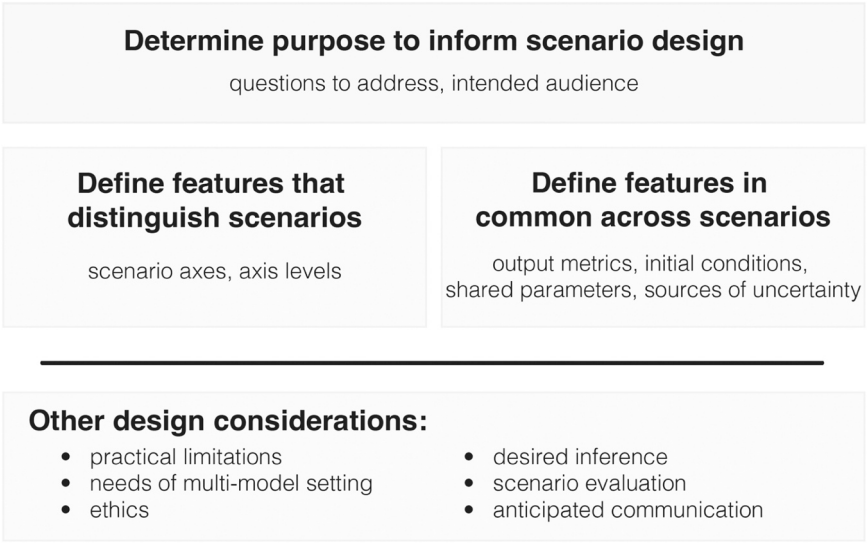

Fig. 2. Overview of scenario design process. First, determine the purpose of the scenario modeling exercise, including the questions to be addressed and the intended audience. This purpose informs all other design decisions. The taxonomy defined in Fig. 1 should be applied at this step. Then, define the features that distinguish scenarios and those that are common across scenarios. Last, consider other design issues that may be relevant during all phases of scenario design.

the parameter of interest and it is common to set values using the associated confidence intervals. For sensitivity analysis or situational awareness designs, values associated with an 80- to 95-percent confidence interval may represent reasonable bounds on current knowledge (Fig. 1(B1) and 1(B2)). Horizon scanning designs may use more extreme values to illustrate what could happen if the future does not conform with the past (Fig. 1(B3)). When empirical confidence intervals are not available, expert opinion, literature review, or survey information can be used to bracket optimistic and pessimistic assumptions (e.g., SMH has used behavioral surveys of propensity to get vaccinated, Beleche et al., 2021). Further considerations on the choice of levels, as related to experimental design, are discussed below.

4.1.3. *Common factors*

Another important aspect of scenario design is the factors that are common across the scenarios. These commonalities are not part of the scenario axes and can include shared data sources and their interpretation, common assumptions about disease dynamics, behavioral responses, or interventions.

4.1.4. *Initial conditions*

Initial conditions represent the state of the modeled system at the start of the scenario projection period. The initial conditions may vary across scenarios or across models. For example, differing scenario assumptions about waning immunity may not only imply a different understanding of what will happen in the future, but also about what occurred in the past; thus, calibration of the model could lead to different initial conditions for each scenario. In a multi-model setting, precise initial conditions are rarely defined, because the models have different calibration approaches and structures. Instead, it is valuable to specify aspects of the process that all models should employ to set initial conditions.

4.1.5. *Additional sources of uncertainty*

While well calibrated forecasts integrate over all sources of uncertainty, scenarios typically encompass a subset of all possible uncertainties (Reich et al., 2022). Forecasting scenario designs represent a special case of scenarios, where the combination of uncertainty captured within a single scenario and between scenarios should be comprehensive (see Fig. 1(B4)). For all other scenario designs, judgment can be made about how much uncertainty to include. The power to discern differences among the scenarios increases as other factors are controlled, but this comes at the expense of generalizability. In a collaborative hub setting, it is important for different modeling teams to make their own choices about many of the uncertainties not specified in the designs. However, it can sometimes be valuable to provide guidance for how to handle key parameters and assumptions that could drive disease dynamics that are not part of the scenario axes (e.g., all SMH rounds provided guidelines on vaccine efficacy, and bounds were often prescribed for waning immunity).

4.2. *Principles of experimental design as applied to scenarios*

As noted above, scenario design and experimental design are closely related conceptually and structurally. Individual scenarios are analogous to experimental treatments, and there are analogous considerations of replication, randomization, and control. Here we briefly discuss these parallels.

The scenario designs proposed in our taxonomy have analogs in experimental design. For example, a 2×2 VOI design is analogous to a randomized block design, where the uncertainty axis serves as a control (or block) variable, to test whether the intervention effect is consistent across blocks (Montgomery, 2017). Designs with more than two axes or levels per axis are also possible, and fractional factorial designs (where only a strategically selected part of the full factorial design is explored) can be used to explore the main effects of many factors, without having to run as many scenarios. Also, there is a tension in statistical design of experiments that helps choose the levels of the factors: the closer the levels are together, the more reasonable it is to assume a linear effect between them; but the farther the levels are apart, the higher the power to discern differences and the scope of inferences that can be made. Similar logic is applicable in scenario design.

In experimental design, holding all factors constant within each replicate is a powerful form of control. Pairing replicates across scenarios is an analogous concept, where as many elements of the model as possible are matched in a particular replicate, like the initial conditions, the sampled parameter values, and, if possible, some aspects of temporal variance. Although common in some fields (e.g., McGowan et al., 2011), pairing replicates can be challenging in epidemiological models, especially if the initial conditions depend on the scenario specification or if demographic stochasticity (e.g., binomial sampling for individual outcomes) is integral to the model (Kaminsky et al., 2019). Nevertheless,

even if it is partial, pairing replicates increases the power to discern treatment effects.

In some cases, multiple rounds of experiments are anticipated, and the results from the early experiments can be used to refine later experiments. Similarly, sequential designs can be achieved with multiple rounds of scenarios. Several SMH rounds were sequential updates in response to the arrival of a new variant or to inform a new policy. For instance, Round 14 was designed to inform the CDC Advisory Committee on Immunization Practices (ACIP) recommendation for reformulated boosters in the fall of 2022, comparing age-restricted versus broader coverage. Presenting results to policymakers (Rosenblum et al., 2022) prompted a follow-up round (Round 15), which made small changes to scenario axes and values to assess whether there would be benefits to releasing boosters earlier (Table 1). If a comparison of outputs across sequential rounds is planned, it is important to record factors that change between rounds, which could confound outputs of interest (e.g., changes in data availability, types of interventions being considered, or new model developments).

More complex scenario designs are possible with multi-round scenarios, including dynamic sequential and branching scenarios. Sequential designs are not fixed *a priori,* but depend on the outcomes of experiments during the exercise (Wald, 1947; Robbins, 1952; Chernoff, 1992). Methods to analyze such sequential statistical designs can be employed to analyze sequential scenarios. Branching scenarios, motivated by branching or nested statistical designs (Hung et al., 2009), can be combined with sequential elements to produce scenarios that are valuable for training and tabletop exercises, where the branches arise in response to dynamic interventions made by users (Barrett et al., 2015; Parikh et al. 2016).

4.3. Other design considerations

4.3.1. Practical limitations

There are practical limitations and trade-offs in scenario design, including model capability, computational resources, clarity of assumptions, and time taken to design a scenario that is actionable. Scenarios must not be too complex, so that modeling teams can generate projections in a reasonable amount of time. In a multi-model setting, minimally complex scenarios also encourage participation from a larger number of teams. Access to additional computational resources can be enhanced in times of crisis, but the need to balance the aims of the scenario design with the practical aspects of modeling remains. If practical constraints strongly and repeatedly influence scenario design, the purposes of the scenario may need to be revisited.

To accommodate these multifaceted needs, the scenario design process is often iterative, involving both internal and external discussions. Internal communication of scenario requirements, especially in the context of a multi-model hub, usually requires multiple rounds of discussion to reduce unwanted (linguistic) uncertainties while retaining a good expression of the scientific uncertainties focal to the scenarios (Shea et al., 2020). If time permits, something akin to the modified Delphi process is valuable: produce a first round of results; discuss the results across models as a group, looking for differences arising from linguistic uncertainty; then allow teams to produce a second round of results that reflect the clarifications (Shea et al., 2020). However, if decision makers only have a short period of time to implement an intervention, a small number of simple scenarios run on stripped-down models might be all that can be achieved.

External discussions with public health decision makers can inform the choice of scenario axes (e.g., potential interventions) and corresponding assumptions (e.g., compliance with those interventions). Curiously, evidence from cognitive psychology suggests that decision makers often need help to fully articulate their concerns (Bond et al., 2008), so a back-and-forth conversation to develop the purpose is an important step. As a result, scenario design can sometimes take several weeks, as illustrated by the SMH experience (Table 1).

4.3.2. Ethical considerations

Ethics of scenario design inherit attributes from the broader ethics of biomedical research (Beauchamp and Childress, 2009), epidemiological research (CIOMS 2009), and decision-making for public health emergencies (Emanuel et al., 2022). Scenario design should have the properties of autonomy, beneficence, non-maleficence, and justice. Autonomy requires the scenario design process to be scientifically grounded and well-documented (including specification of clear objectives, Smith et al., 2021), capturing uncertainty and recording assumptions. Additionally, scenario design should be beneficial, in that it should promote evidence-based policymaking (Choi et al., 2005), providing benefit over decisions that would be made without scenario projections (Taylor, 2003). These principles also play a role in non-maleficence, as inaccuracies in scenario design may cause harm to populations affected by the recommendations. Simultaneously, affected populations may have conflicting perspectives on the decision alternatives (Brownson et al., 2013), which can ideally be incorporated into scenario design via counterfactual null scenarios or multi-objective formulations. Existing inequalities should be incorporated into and addressed by scenario design; further, scenario design should not exacerbate these inequalities, nor create new inequalities (CIOMS 2009). Scenario modeling efforts should be evaluated according to these criteria (Boden and McKendrick, 2017). Multiple metrics of equity and fairness can be considered in scenario design (Braveman and Gruskin, 2003; Whitehead, 1992; Mhasawade et al., 2021), as different stakeholders may have different perspectives (Whitehead, 1992).

4.3.3. Scenario evaluation

In some cases, there may be a desire to evaluate scenarios and projections after the projection period has passed; does this desire affect scenario design? Broadly, scenarios are well designed if the resulting projections answer the primary question and serve the intended users even if the scenario assumptions do not materialize. Yet it still may be useful to assess how well scenario assumptions match unfolding reality, especially when a goal of scenario design is bracketing (i.e., situational awareness or forecasting situations). Scenario evaluation is difficult in practice; Howerton et al. (2023a) provide an illustration of salient issues. Scenario parameters may not be measurable even after projection periods have passed (e.g., degree of immune escape of a new variant, or even the impact of an NPI). For horizon scanning and for scenarios including counterfactuals, evaluation of scenario parameters and resulting projections can be particularly difficult. But against these challenges, there can be meaningful benefits of scenario evaluation, including ongoing improvement of individual models (Alley et al., 2019), improvement in scenario design itself, and building trust with end-users (Raftery, 2016). Consideration of scenario evaluation at the design stage may enhance the ability to evaluate the scenarios, for instance, by choosing scenario axes that can eventually be measured. Thus, while scenario evaluation is not the primary purpose of a scenario projection exercise, it might be a strategic consideration in scenario design.

4.3.4. Communication

In epidemiology, scenario projections are often designed for specific audiences and can have public visibility. Thus, it can be advantageous for scenarios to be clear enough for easy communication to and interpretation by external audiences. For example, SMH uses a standard scenario design template to provide consistency in how scenario assumptions are shared and ease comparisons between rounds and between hubs (see https://github.com/midas-network/covid19-scenario-modeling-hub for an example). For additional discussion of generalizable infrastructure, see the HubVerse project (https://hubdocs.readthedocs.io/).

Presentation of scenario results is also an important component of communication worth anticipating. A successful display of results entails three, often conflicting, objectives: enabling comparisons among

scenarios; communicating the uncertainty within and across models (in a multi-model setting); and supporting multiple different classes of constituents, including researchers, public health officials, journalists, and members of the public. Key challenges include communication of the nuances of scenario projections to lay audiences (as different from forecasts), and visualization of uncertainty (Kamal et al., 2021; Hullman et al., 2019; Hägele et al., 2022; Spiegelhalter, 2017). See Loo et al. (2024) for further discussion of SMH communication strategies.

5. Discussion

The experience of the U.S. COVID-19 Scenario Modeling Hub over its first 17 rounds provided an impetus for the scenario taxonomy proposed in this work, which we believe will be valuable in epidemiological settings, and perhaps more broadly. We have attempted to provide broad guidelines for scenario design that apply in single and multi-model efforts, and made parallels with other fields such as experimental design. Several insights with broader relevance bear reflection: the importance of the audience and a clear statement of the purpose of the design; the power of the design itself; the need to think carefully about uncertainty; and the benefits of a clear process.

By projecting multiple, clearly defined scenarios that were motivated by public health needs, SMH projections have had significant public health impact (Borchering et al., 2021, 2023; Truelove et al., 2022; Rosenblum et al., 2022; Biggerstaff et al., 2022). It was difficult at times, however, to balance the needs of decision makers with the capabilities of available models. Implementing realistic scenarios and generating well-calibrated projections can require added model complexity or additional time. Key policy questions or vast uncertainty may suggest the need for many scenarios, but computational constraints may limit the number of scenarios that can be modeled in a timely fashion. Further, the foundational philosophy behind multi-model ensembles, namely, the diversity of approaches taken by the independent groups (Shea et al., 2020) can itself pose a challenge for scenario design. But the repeated nature of the SMH effort has allowed the complexity of the models and the subtleties of scenarios to increase.

Clarity of audience and purpose affect scenario design and its impact. The influence of SMH rounds that were developed in direct conversation with a decision-making agency was easiest to illustrate. But many valuable impacts are harder to demonstrate, like the deepening understanding among modeling teams of the epidemic in the U.S. prompted by the structured challenge of shared scenarios. In retrospect, we found it somewhat challenging to look back over two years of work and recover the specific purposes of each round. We propose that an active and clear articulation of the audiences and the purposes of a scenario design will help to sharpen the design of the scenarios in future SMH rounds, and similar efforts in the future.

The design of scenarios provides the structure for inference. The taxonomy captured in Fig. 1 was not available during the design of SMH Rounds 1 through 17, but we believe that it could have enhanced some of the designs. Value of information designs are particularly interesting, because they both allow the comparison of alternative interventions and test those comparisons against critical sources of uncertainty. The SMH used VOI designs in 6 of the first 17 rounds, but curiously, none of those showed a reversal in preference of intervention based on the uncertainty axis. On one hand, that's a great relief to decision makers, but on the other hand, it raises a question about whether the uncertainties considered were most relevant to the decisions. Would conscious attention to the power of particular design structures lead to even more valuable scenario designs?

Scenario design invites careful and deliberate consideration of uncertainties. Scenario axes often focus on uncertainties hypothesized to be major drivers of future dynamics or decision outcomes. The first question is whether the process used to identify those uncertainties is robust. The second, perhaps more difficult question, is how to handle the remaining uncertainty. For example, operational uncertainties about the implementation of interventions may be required to create clear, easy-to-interpret scenarios, but such uncertainties are also important to account for in projected outcomes. In evaluating outcomes of the early SMH rounds, Howerton et al. (2023a) and Wade-Malone et al. (2024) note that results of individual models often had quite different variances, suggesting that they captured different sets of uncertainty that weren't otherwise specified in the scenario design. Is that problematic or desirable? How does calibration of the individual models affect calibration of the ensemble projection (Howerton et al., 2023b), and how does that affect scenario design? We believe that there are some open questions here that warrant further study.

The process of scenario design affects efficiency, participation, trust, and communication. Particularly in multi-model collaborative settings, the process of scenario design is challenging, and a clear process with dedicated support staff can support and invite the participation of the collaborating teams. But even in single-model settings, the process of scenario design aids in communication with the intended audiences and can promote trust.

As noted earlier, scenario design is practiced in many fields besides epidemiology. The framework that we have proposed in this paper integrates elements of the three schools described by Amer et al. (2013): like the Intuitive Logics School, it relies on experts' conceptual understanding of systems to develop causal maps that inform scenario design; like the Probabilistic Modified Trends School, it combines extrapolation of past trends with modifications to acknowledge changes in the future; and, like the French School, it places an emphasis on the decision setting, that is, the ways in which trajectories can be influenced by intervention. We are hopeful that more cross-disciplinary examination of how scenario projections are designed and used can lead to common advances across fields.

Beyond the insights that arose from the SMH experience, there are other questions that may be relevant for development of scenario modeling practices in epidemiology. We see these as open questions for a future research agenda, to improve the impact of scenario methods:

- Are there perspectives in scenario design, especially in collaborative ensemble settings in other fields, that would enhance the practice in epidemiology? Similarly, can our efforts inform practice in other fields?
- Is there a design trajectory across an epidemic? That is, can we anticipate a specific sequence of questions, and even have template scenario designs ready? As a proposal, four stages could arise: (1) initial bounding of uncertainty and exploration of simple interventions; (2) assessing specific interventions (e.g., vaccination) as they become available; (3) assessing new dynamics (e.g., variants, behavior changes) as they arise; and (4) transitions to questions relevant in an endemic phase. Is this a useful start?
- What are the pitfalls to avoid in scenario design in public health settings? Is it possible to inadvertently mislead or confuse decision makers with poorly designed scenarios? Are other unintentional negative outcomes possible?
- Were there an operational scenario modeling hub for a particular disease, would a set of scenario designs become standard? For instance, would a regularly calibrated baseline scenario with several updated contrasts (e.g., emergence of a new variant) make sense? Or are infectious diseases too complex, because human behavior, available interventions, and viral evolution change so quickly that standard scenario designs are not useful?
- Can lessons from scenario design in one location or outbreak reliably be applied in other settings? This will be particularly important at times where urgent results are needed, or in low-resource settings (e.g., low and middle income countries).
- How can we best communicate scenario results and explain the difference between scenarios and forecasts, which are more intuitive? Are verbal or numerical or graphical representations of scenario

designs and results most effective and do they differ for more or less quantitatively comfortable users?

In summary, the COVID-19 pandemic coalesced a great deal of burgeoning expertise in epidemiological modeling, scenario projection, scenario design, and collaborative modeling endeavors. Using the experience of the U.S. Scenario Modeling Hub to reflect on the state-of-the-art in scenario design, we believe that a sound philosophical framework and procedural methodology for scenario design would increase the efficiency and efficacy of these methods, both in epidemiological settings and in other fields of endeavor.

CRediT authorship contribution statement

Shaun Truelove: Writing – review & editing. **Harry Hochheiser:** Writing – original draft, Writing – review & editing. **Katie Yan:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Cécile Viboud:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Emily Howerton:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Justin Lessler:** Writing – review & editing. **Katriona Shea:** Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Madhav V Marathe:** Investigation, Writing – original draft, Writing – review & editing. **Rebecca Borchering:** Writing – review & editing. **William J.M. Probert:** Investigation, Writing – review & editing. **Erik Rosenstrom:** Investigation, Writing – original draft, Writing – review & editing. **Michael C Runge:** Conceptualization, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Srinivasan Venkatramanan:** Writing – review & editing. **Bryan Lewis:** Writing – review & editing.

Declaration of Competing Interest

MCR reports stock ownership in Becton Dickinson & Co., which manufactures medical equipment used in COVID-19 testing, vaccination, and treatment. JL has served as an expert witness on cases where the likely length of the pandemic was of issue. There are no other competing interests to declare.

Acknowledgments

We are grateful to all the participants in the U.S. COVID-19 Scenario Modeling Hub, for stimulating conversations over the last two-and-a-half years that have motivated the ideas in this paper. KS, EH, and RB acknowledge funding from NSF COVID-19 RAPID awards DEB-2028301, DEB-2037885, DEB-2126278, and DEB-2220903. EH was supported by the Eberly College of Science Barbara McClintock Science Achievement Graduate Scholarship in Biology at the Pennsylvania State University. KY was supported by the Penn State Verne M. Willaman Distinguished Fellowship in Science and NSF Grant No. DGE1255832. HH was supported by NIGMS Grant U24GM132013. MM, SV, and BL acknowledge support from NSF Expeditions in Computing Grant CCF-1918656, VDH Contract UVABIO610-GY23, and DTRA/ARA contract S-D00189-IDIQ-15-UVA.

Disclaimers

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the US National Institutes of Health or Department of Health and Human Services. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Alley, R.B., Emanuel, K.A., Zhang, F., 2019. Advances in weather prediction. Science 363, 342–344. https://doi.org/10.1126/science.aav7274.

Amer, M., Daim, T.U., Jetter, A., 2013. A review of scenario planning. Futures 46, 23–40. https://doi.org/10.1016/j.futures.2012.10.003.

Bartholomew, K., 2007. Land use-transportation scenario planning: promise and reality. Transportation 34, 397–412. https://doi.org/10.1007/s11116-006-9108-2.

Barrett, C., Eubank, S., Marathe, A., Marathe, M., Swarup, S., 2015. Synthetic information environments for policy informatics: a distributed cognition perspective. In: Johnston, E.W. (Ed.), Governance in the Information Era: Theory and Practice of Policy Informatics. Routledge, New York, NY, pp. 267–284.

Bay, C., St-Onge, G., Davis, J.T., Chinazzi, M., Howerton, E., Lessler, J., Runge, M.C., Shea, K., Truelove, S., Viboud, C., Vespigiani, A., 2024. Ensemble^2: scenarios ensembling for communication and performance analysis. Epidemics 46, 100748. https://doi.org/10.1016/j.epidem.2024.100748.

Beauchamp, T.L., Childress, J.F., 2009. Principles of biomedical ethics, 6th ed. Oxford University Press New York, New York. ISBN 9780195335705.

Beleche, T., Ruhter, J., Kolbe, A., Marus, J., Bush, L., Sommers, B., 2021. COVID-19 vaccine hesitancy: demographic factors, geographic patterns, and changes over time. Issue Brief, May 2021. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, Washington, DC. 27 p. ⟨https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//200816/aspe-ib-vaccine-hesitancy.pdf⟩.

Biggerstaff, M., Slayton, R.B., Johansson, M.A., Butler, J.C., 2022. Improving pandemic response: employing mathematical modeling to confront coronavirus disease 2019. Clin. Infect. Dis. 74, 913–917. https://doi.org/10.1093/cid/ciab673.

Boden, L.A., McKendrick, I.J., 2017. Model-based policymaking: a framework to promote ethical "good practice" in mathematical modeling for public health policymaking. Front. Public Health 5, 68. https://doi.org/10.3389/fpubh.2017.00068.

Bond, S.D., Carlson, K.A., Keeney, R.L., 2008. Generating objectives: can decision makers articulate what they want? Manag. Sci. 54, 56–70. https://doi.org/10.1287/mnsc.1070.0754.

Borchering, R.K., Mullany, L.C., Howerton, E., Chinazzi, M., Smith, C.P., Qin, M., Reich, N.G., Contamin, L., Levander, J., Kerr, J., Espino, J., Hochheiser, H., Lovett, K., Kinsey, M., Tallaksen, K., Wilson, S., Shin, L., Lemaitre, J.C., Hulse, J.D., Kaminsky, J., Lee, E.C., Hill, A.L., Davis, J.T., Mu, K., Xiong, X., Pastore y Piontti, A., Vespignani, A., Srivastava, A., Porebski, P., Venkatramanan, S., Adiga, A., Lewis, B., Klahn, B., Outten, J., Hurt, B., Chen, J., Mortveit, H., Wilson, A., Marathe, M., Hoops, S., Bhattacharya, P., Machi, D., Chen, S., Paul, R., Janies, D., Thill, J.-C., Galanti, M., Yamana, T., Pei, S., Shaman, J., España, G., Cavany, S., Moore, S., Perkins, A., Healy, J.M., Slayton, R.B., Johansson, M.A., Biggerstaff, M., Shea, K., Truelove, S.A., Runge, M.C., Viboud, C., Lessler, J., 2023. Impact of SARS-CoV-2 vaccination of children ages 5–11 years on COVID-19 disease burden and resilience to new variants in the United States, November 2021–March 2022: a multi-model study. Lancet Reg. Health - Am. 17, 100398 https://doi.org/10.1016/j.lana.2022.100398.

Borchering, R.K., Viboud, C., Howerton, E., Smith, C.P., Truelove, S., Runge, M.C., Reich, N.G., Contamin, L., Levander, J., Salerno, J., others, 2021. Modeling of future COVID-19 cases, hospitalizations, and deaths, by vaccination rates and nonpharmaceutical intervention scenarios—United States, April–September 2021. Morb. Mortal. Wkly. Rep. 70, 719. https://doi.org/10.15585/mmwr.mm7019e3.

Braveman, P., Gruskin, S., 2003. Defining equity in health. J. Epidemiol. Community Health 57, 254–258. https://doi.org/10.1136/jech.57.4.254.

Brownson, R.C., Fielding, J.E., Maylahn, C.M., 2013. Evidence-based decision making to improve public health practice. Front Public Health Serv. Syst. Res 2, 2. https://doi.org/10.13023/FPHSSR.0202.02.

Chadès, I., Carwardine, J., Martin, T.G., Nicol, S., Sabbadin, R., Buffet, O., 2012. MOMDPs: a solution for modelling adaptive management problems. Proc. AAAI Conf. Artif. Intell. 26, 267–273. https://doi.org/10.1609/aaai.v26i1.8171.

Chernoff, H., 1992. Sequential design of experiments. In: Kotz, S., Johnson, N.L. (Eds.), Breakthroughs in Statistics. Springer, New York, New York, pp. 345–360. https://doi.org/10.1007/978-1-4612-4380-9_27.

Choi, B.C., Pang, T., Lin, V., Puska, P., Sherman, G., Goddard, M., Ackland, M.J., Sainsbury, P., Stachenko, S., Morrison, H., 2005. Can scientists and policy makers work together? J. Epidemiol. Community Health 59, 632–637. https://doi.org/10.1136/jech.2004.031765.

Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.-J., Vespignani, A., 2007. Modeling the Worldwide Spread of Pandemic Influenza: baseline Case and Containment Interventions. PLOS Med. 4, e13 https://doi.org/10.1371/journal.pmed.0040013.

Cordova-Pozo, K., Rouwette, E.A., 2023. Types of scenario planning and their effectiveness: A review of reviews. Futures 149, 103153. https://doi.org/10.1016/j.futures.2023.103153.

[CIOMS] Council for International Organizations of Medical Sciences, 2009. International ethical guidelines on epidemiological studies. World Health Organization. 127 p. ISBN 9789290360810.

Cramer, E.Y., Ray, E.L., Lopez, V.K., Bracher, J., Brennen, A., Castro Rivadeneira, A.J., Gerding, A., Gneiting, T., House, K.H., Huang, Y., others, 2022. Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States. Proc. Natl. Acad. Sci. 119, e2113561119 https://doi.org/10.1073/pnas.2113561119.

Dausey, D.J., Lurie, N., Diamond, A., Meade, B., Molander, R., Ricci, K., Stoto, M., Wasserman, J., 2005. Bioterrorism Preparedness Training and Assessment Exercises for Local Public Health Agencies (No. TR-261-DHHS). RAND Corporation, Santa Monica, CA. ⟨https://www.rand.org/content/dam/rand/pubs/technical_reports/2005/RAND_TR261.pdf⟩.

DEFRA] Department for Environment Food & Rural Affairs, 2018. Exercise Blackthorn: national foot and mouth disease exercise evaluation and lessons identified report. Animal and Plant Health Agency, London, United Kingdom, p. 67. doi: ⟨https://www.gov.uk/government/publications/foot-and-mouth-disease-exercise-blackthorn-evaluation-2018⟩.

Dowse, A., 2021. Scenario planning methodology for future conflict. J. Indo-Pac. Aff. 4 (2), 18–31. ⟨https://media.defense.gov/2021/Mar/07/2002595034/-1/-1/1/DOWSE.PDF⟩.

Eaton, J.W., Johnson, L.F., Salomon, J.A., Bärnighausen, T., Bendavid, E., Bershteyn, A., Bloom, D.E., Cambiano, V., Fraser, C., Hontelez, J.A., 2012. HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med. 9, e1001245 https://doi.org/10.1371/journal.pmed.1001245.

Emanuel, E.J., Upshur, R.E., Smith, M.J., 2022. What Covid has taught the world about ethics. N. Engl. J. Med. 387, 1542–1545. https://doi.org/10.1056/NEJMp2210173.

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. 9, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016.

Felli, J.C., Hazen, G.B., 1998. Sensitivity analysis and the expected value of perfect information. Med. Decis. Mak. 18, 95–109. https://doi.org/10.1177/0272989X9801800117.

Flasche, S., Jit, M., Rodríguez-Barraquer, I., Coudeville, L., Recker, M., Koelle, K., Milne, G., Hladish, T.J., Perkins, T.A., Cummings, D.A., 2016. The long-term safety, public health impact, and cost-effectiveness of routine vaccination with a recombinant, live-attenuated dengue vaccine (Dengvaxia): a model comparison study. PLoS Med. 13, e1002181 https://doi.org/10.1371/journal.pmed.1002181.

[FDA] Food and Drug Administration, 2020. FDA takes additional action in fight against COVID-19 by issuing emergency use authorization for second COVID-19 vaccine. News Release, December 18, 2020. ⟨https://www.fda.gov/news-events/press-announcements/fda-takes-additional-action-fight-against-covid-19-issuing-emergency-use-authorization-second-covid⟩.

Hägele, D., Schulz, C., Beschle, C., Booth, H., Butt, M., Barth, A., Deussen, O., Weiskopf, D., 2022. Uncertainty visualization: Fundamentals and recent developments. it-Inf. Technol. 64, 121–132. https://doi.org/10.1515/itit-2022-0033.

Hellewell, J., Abbott, S., Gimma, A., Bosse, N.I., Jarvis, C.I., Russell, T.W., Munday, J.D., Kucharski, A.J., Edmunds, W.J., Sun, F., Flasche, S., Quilty, B.J., Davies, N., Liu, Y., Clifford, S., Klepac, P., Jit, M., Diamond, C., Gibbs, H., van Zandvoort, K., Funk, S., Eggo, R.M., 2020. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob. Health 8, e488–e496. https://doi.org/10.1016/S2214-109X(20)30074-7.

Henderson, D.A., Inglesby Jr., T.V., O'Toole, T., Inglesby, T.V., Grossman, R., O'Toole, T., 2001. A Plague on Your City: observations from TOPOFF. Clin. Infect. Dis. 32, 436–445. https://doi.org/10.1086/318513.

Houben, R.M., Menzies, N.A., Sumner, T., Huynh, G.H., Arinaminpathy, N., Goldhaber-Fiebert, J.D., Lin, H.-H., Wu, C.-Y., Mandal, S., Pandey, S., 2016. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. Lancet Glob. Health 4, e806–e815. https://doi.org/10.1016/S2214-109X(16)30199-1.

Howerton, E., Contamin, L., Mullany, L.C., Qin, M., Reich, N.G., Bents, S., Borchering, R. K., Jung, S.-m, Loo, S.L., Smith, C.P., Levander, J., Kerr, J., Espino, J., Panhuis, W. Gv, Hochheiser, H., Galanti, M., Yamana, T., Pei, S., Shaman, J., Rainwater-Lovett, K., Kinsey, M., Tallaksen, K., Wilson, S., Shin, L., Lemaitre, J.C., Kaminsky, J., Hulse, J.D., Lee, E.C., McKee, C., Hill, A., Karlen, D., Chinazzi, M., Davis, J.T., Mu, K., Xiong, X., Piontti, A.Py, Vespignani, A., Rosenstrom, E.T., Ivy, J.S., Mayorga, M.E., Swann, J.L., España, G., Cavany, S., Moore, S., Perkins, A., Hladish, T., Pillai, A., Toh, K.B., Ira Longini, J., Chen, S., Paul, R., Janies, D., Thill, J.-C., Bouchnita, A., Bi, K., Lachmann, M., Fox, S., Meyers, L.A., Consortium, U.C.-M., Srivastava, A., Porebski, P., Venkatramanan, S., Adiga, A., Lewis, B., Klahn, B., Outten, J., Hurt, B., Chen, J., Mortveit, H., Wilson, A., Marathe, M., Hoops, S., Bhattacharya, P., Machi, D., Cadwell, B.L., Healy, J.M., Slayton, R.B., Johansson, M.A., Biggerstaff, M., Truelove, S., Runge, M.C., Shea, K., Viboud, C., Lessler, J., 2023a. Informing pandemic response in the face of uncertainty: an evaluation of the U.S. COVID-19 Scenario Modeling Hub. Nature Communications. ⟨https://doi.org/10.1038/s41467-023-42680-x⟩.

Howerton, E., Runge, M.C., Bogich, T.L., Borchering, R.K., Inamine, H., Lessler, J., Mullany, L.C., Probert, W.J., Smith, C.P., Truelove, S., 2023b. Context-dependent representation of within-and between-model uncertainty: aggregating probabilistic predictions in infectious disease epidemiology. J. R. Soc. Interface 20, 20220659. https://doi.org/10.1098/rsif.2022.0659.

Hullman, J., Qiao, X., Correll, M., Kale, A., Kay, M., 2019. In pursuit of error: A survey of uncertainty visualization evaluation. IEEE Trans. Vis. Comput. Graph. 25, 903–913. https://doi.org/10.1109/TVCG.2018.2864889.

Hung, Y., Joseph, V.R., Melkote, S.N., 2009. Design and analysis of computer experiments with branching and nested factors. Technometrics 51, 354–365. https://doi.org/10.1198/TECH.2009.07097.

Johansson, M.A., Apfeldorf, K.M., Dobson, S., Devita, J., Buczak, A.L., Baugher, B., Moniz, L.J., Bagley, T., Babin, S.M., Guven, E., others, 2019. An open challenge to advance probabilistic forecasting for dengue epidemics. Proc. Natl. Acad. Sci. 116, 24268–24274. https://doi.org/10.1073/pnas.1909865116.

Johnson, F.A., Moore, C.T., Kendall, W.L., Dubovsky, J.A., Caithamer, D.F., Kelley Jr., J. R., Williams, B.K., 1997. Uncertainty and the management of mallard harvests. J. Wildl. Manag. 61, 202–216. https://doi.org/10.2307/3802429.

Juul, J.L., Græsbøll, K., Christiansen, L.E., Lehmann, S., 2021. Fixed-time descriptive statistics underestimate extremes of epidemic curve ensembles. Nat. Phys. 17, 5–8. https://doi.org/10.1038/s41567-020-01121-y.

Kamal, A., Dhakal, P., Javaid, A.Y., Devabhaktuni, V.K., Kaur, D., Zaientz, J., Marinier, R., 2021. Recent advances and challenges in uncertainty visualization: a survey. J. Vis. 24, 861–890. https://doi.org/10.1007/s12650-021-00755-1.

Kaminsky, J., Keegan, L.T., Metcalf, C.J.E., Lessler, J., 2019. Perfect counterfactuals for epidemic simulations. Philos. Trans. R. Soc. B: Biol. Sci. 374, 20180279. https://doi.org/10.1098/rstb.2018.0279.

Keeney, R.L., Raiffa, H., 1993. Decisions with multiple objectives: preferences and value tradeoffs. Cambridge University Press, Cambridge, United Kingdom. ISBN 0521438837.

Khakee, A., 1991. Scenario construction for urban planning. Omega 19, 459–469. https://doi.org/10.1016/0305-0483(91)90062-X.

Kim, K., Park, T., Pastrana, J., Marin, M., Cortes, E.A., Rabelo, L.C., Lee, G., 2014. Modeling of complex scenarios using LVC simulation. Proceedings of the 2014 Winter Simulation Conference. IEEE Press, Savannah, Georgia, pp. 2931–2941. https://dl.acm.org/doi/abs/10.5555/2693848.2694219.

Krey, V., 2014. Global energy-climate scenarios and models: a review. Wiley Interdiscip. Rev.: Energy Environ. 3, 363–383. https://doi.org/10.1002/wene.98.

Leung, D.Y., Yang, Y., 2012. Wind energy development and its environmental impact: a review. Renew. Sustain. Energy Rev. 16, 1031–1039. https://doi.org/10.1016/j.rser.2011.09.024.

Longini Jr, I.M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D.A., Halloran, M.E., 2005. Containing pandemic influenza at the source. Science 309, 1083–1087. https://doi.org/10.1126/science.1115717.

Loo, S., Howerton, E., Contamin, L., Smith, C.P., Borchering, R.K., Mullany, L.C., Bents, S., Carcelen, E., Jung, S., Bogich, T., van Panhuis, W.G., Kerr, J., Espino, J., Yan, K., Hochheiser, H., Runge, M.C., Shea, K., Lessler, J., Viboud, C., Truelove, S., 2024. The US COVID-19 and Influenza Scenario Modeling Hubs: delivering long-term projections to guide policy. Epidemics 46, 100738. https://doi.org/10.1016/j.epidem.2023.100738.

Martelli, A., 2014. Models of scenario building and planning: Facing uncertainty and complexity. Palgrave Macmillan, United Kingdom. ⟨https://doi.org/10.1057/9781137293503⟩.

McDowall, W., Eames, M., 2006. Forecasts, scenarios, visions, backcasts and roadmaps to the hydrogen economy: A review of the hydrogen futures literature. Energy Policy 34, 1236–1250. https://doi.org/10.1016/j.enpol.2005.12.006.

McGowan, C.P., Runge, M.C., Larson, M.A., 2011. Incorporating parametric uncertainty into population viability analysis models. Biol. Conserv. 144, 1400–1408. https://doi.org/10.1016/j.biocon.2011.01.005.

Meehl, G.A., Boer, G.J., Covey, C., Latif, M., Stouffer, R.J., 2000. The coupled model intercomparison project (CMIP). Bull. Am. Meteorol. Soc. 81, 313–318. ⟨https://www.jstor.org/stable/26215108⟩.

Meltzer, M.I., Atkins, C.Y., Santibanez, S., Knust, B., Petersen, B.W., Ervin, E.D., Nichol, S.T., Damon, I.K., Washington, M.L., 2014. Estimating the future number of cases in the Ebola epidemic–Liberia and Sierra Leone, 2014–2015. MMWR Suppl. 63 (3), 1–14. ⟨https://stacks.cdc.gov/view/cdc/24901⟩.

Mhasawade, V., Zhao, Y., Chunara, R., 2021. Machine learning and algorithmic fairness in public and population health. Nat. Mach. Intell. 3, 659–666. https://doi.org/10.1038/s42256-021-00373-4.

Montgomery, D.C., 2017. Design and analysis of experiments, Tenth ed. John Wiley & Sons, Inc., Hoboken, New Jersey.

NASEM] National Academies of Science, Engineering, and Medicine, 2023. Risk analysis methods for nuclear war and nuclear terrorism. 117 p. The National Academies Press, Washington, DC. ⟨https://doi.org/10.17226/26609⟩.

[NRC] National Research Council, 2008. Behavioral Modeling and Simulation: From Individuals to Societies. The National Academies Press, Washington, DC. doi:10.17226/12169.

Nicholson, E., Fulton, E.A., Brooks, T.M., Blanchard, R., Leadley, P., Metzger, J.P., Mokany, K., Stevenson, S., Wintle, B.A., Woolley, S.N., 2019. Scenarios and models to support global conservation targets. Trends Ecol. Evol. 34, 57–68. https://doi.org/10.1016/j.tree.2018.10.006.

Parikh, N., Hayatnagarkar, H.G., Beckman, R.J., Marathe, M.V., Swarup, S., 2016. A comparison of multiple behavior models in a simulation of the aftermath of an improvised nuclear detonation. Auton. Agents Multi-Agent Syst. 30, 1148–1174. https://doi.org/10.1007/s10458-016-9331-y.

Pfister, S., Bayer, P., Koehler, A., Hellweg, S., 2011. Projected water consumption in future global agriculture: scenarios and related impacts. Sci. Total Environ. 409, 4206–4216. https://doi.org/10.1016/j.scitotenv.2011.07.019.

Pitzer, V.E., Viboud, C., Simonsen, L., Steiner, C., Panozzo, C.A., Alonso, W.J., Miller, M. A., Glass, R.I., Glasser, J.W., Parashar, U.D., 2009. Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. Science 325, 290–294. https://doi.org/10.1126/science.1172330.

Prasad, P.V., Steele, M.K., Reed, C., Meyers, L.A., Du, Z., Pasco, R., Alfaro-Murillo, J.A., Lewis, B., Venkatramanan, S., Schlitt, J., Chen, J., Orr, M., Wilson, M.L., Eubank, S., Wang, L., Chinazzi, M., Pastore y Piontti, A., Davis, J.T., Halloran, M.E., Longini, I., Vespignani, A., Pei, S., Galanti, M., Kandula, S., Shaman, J., Haw, D.J., Arinaminpathy, N., Biggerstaff, M., 2023. Multimodeling approach to evaluating the efficacy of layering pharmaceutical and nonpharmaceutical interventions for influenza pandemics. Proc. Natl. Acad. Sci. 120, e2300590120 https://doi.org/10.1073/pnas.2300590120.

Probert, W.J.M., Shea, K., Fonnesbeck, C.J., Runge, M.C., Carpenter, T.E., Dürr, S., Garner, M.G., Harvey, N., Stevenson, M.A., Webb, C.T., Werkman, M., Tildesley, M. J., Ferrari, M.J., 2016. Decision-making for foot-and-mouth disease control:

Objectives matter. Epidemics 15, 10–19. https://doi.org/10.1016/j.epidem.2015.11.002.
Raftery, A.E., 2016. Use and communication of probabilistic forecasts. Stat. Anal. Data Min.: ASA Data Sci. J. 9, 397–410. https://doi.org/10.1002/sam.11302.
Reich, N.G., Lessler, J., Funk, S., Viboud, C., Vespignani, A., Tibshirani, R.J., Shea, K., Schienle, M., Runge, M.C., Rosenfeld, R., Ray, E.L., Niehus, R., Johnson, H.C., Johansson, M.A., Hochheiser, H., Gardner, L., Bracher, J., Borchering, R.K., Biggerstaff, M., 2022. Collaborative hubs: making the most of predictive epidemic modeling. Am. J. Public Health 112, 839–842. https://doi.org/10.2105/AJPH.2022.306831.
Robbins, H., 1952. Some aspects of the sequential design of experiments. Bull. Am. Math. Soc. 58, 527–535. https://doi.org/10.1090/S0002-9904-1952-09620-8.
Rosenblum, H.G., Wallace, M., Godfrey, M., Roper, L.E., Hall, E., Fleming-Dutra, K.E., Link-Gelles, R., Pilishvili, T., Williams, J., Moulia, D.L., Brooks, O., Talbot, H.K., Lee, G.M., Bell, B.P., Daley, M.F., Meyer, S., Oliver, S.E., Twentyman, E., 2022. Interim recommendations from the Advisory Committee on Immunization Practices for the use of bivalent booster doses of COVID-19 vaccines—United States, October 2022. Morb. Mortal. Wkly. Rep. 71, 1436–1441. https://doi.org/10.15585/mmwr.mm7145a2.
Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity analysis in practice: a guide to assessing scientific models. 232 p. John Wiley & Sons, West Sussex, England. ISBN 0470870931.
Shafer, G., 1990. Perspectives on the theory and practice of belief functions. Int. J. Approx. Reason. 4, 323–362. https://doi.org/10.1016/0888-613X(90)90012-Q.
Shea, K., Borchering, R.K., Probert, W.J., Howerton, E., Bogich, T.L., Li, S.-L., van Panhuis, W.G., Viboud, C., Aguás, R., Belov, A.A., Bhargava, S.H., Cavany, S., Chang, J.C., Chen, C., Chen, J., Chen, S., Chen, Y., Childs, L.M., Chow, C.C., Crooker, I., Del Valle, S.Y., España, G., Fairchild, G., Gerkin, R.C., Germann, T.C., Gu, Q., Guan, X., Guo, L., Hart, G.R., Hladish, T.J., Hupert, N., Janies, D., Kerr, C.C., Klein, D.J., Klein, E., Lin, G., Manore, C., Meyers, L.A., Mittler, J., Mu, K., Núñez, R.C., Oidtman, R., Pasco, R., Pastore y Piontti, A., Paul, R., Pearson, C.A.B., Perdomo, D.R., Perkins, T.A., Pierce, K., Pillai, A.N., Rael, R.C., Rosenfeld, K., Ross, C.W., Spencer, J.A., Stoltzfus, A.B., Toh, K.B., Vattikuti, S., Vespignani, A., Wang, L., White, L., Xu, P., Yang, Y., Yogurtcu, O.N., Zhang, W., Zhao, Y., Zou, D., Ferrari, M.J., Pannell, D., Tildesley, M.J., Seifarth, J., Johnson, E., Biggerstaff, M., Johansson, M., Slayton, R.B., Levander, J., Stazer, J., Salerno, J., Runge, M.C., 2023. Multiple models for outbreak decision support in the face of uncertainty. Proc. Natl. Acad. Sci. 120, e2207537120 https://doi.org/10.1073/pnas.2207537120.
Shea, K., Kelly, D., Sheppard, A.W., Woodburn, T.L., 2005. Context-dependent biological control of an invasive thistle. Ecology 86, 3174–3181. https://doi.org/10.1890/05-0195.
Shea, K., Runge, M.C., Pannell, D., Probert, W.J., Li, S.-L., Tildesley, M., Ferrari, M., 2020. Harnessing multiple models for outbreak management. Science 368, 577–579. https://doi.org/10.1126/science.abb9934.
Sherratt, K., Srivastava, A., Ainslie, K., Singh, D.E., Cublier, A., Marinescu, M.C., Carretero, J., Garcia, A.C., Franco, N., Willem, L., Abrams, S., Faes, C., Beutels, P., Hens, N., Müller, S., Charlton, B., Ewert, R., Paltra, S., Rakow, C., Rehmann, J., Conrad, T., Schütte, C., Nagel, K., Abbott, S., Grah, R., Niehus, R., Prasse, B., Sandmann, F., Funk, S., 2024. Characterising information gains and losses when collecting multiple epidemic model outputs. Epidemics 46, 100765. https://doi.org/10.1016/j.epidem.2024.100765.
Smith, M.J., Ahmad, A., Arawi, T., Dawson, A., Emanuel, E.J., Garani-Papadatos, T., Ghimire, P., Iliyasu, Z., Lei, R., Mastroleo, I., Mathur, R., Okeibunor, J., Parker, M., Saenz, C., Thomé, B., Upshur, R.E.G., Voo, T.C., 2021. Top five ethical lessons of COVID-19 that the world must learn. Wellcome Open Res. 6, 17. https://doi.org/10.12688/wellcomeopenres.16568.1.
Spaniol, M.J., Rowland, N.J., 2019. Defining scenario. Futures Foresight Sci. 1, e3 https://doi.org/10.1002/ffo2.3.
Spiegelhalter, D., 2017. Risk and uncertainty communication. Annu. Rev. Stat. Appl. 4, 31–60. https://doi.org/10.1146/annurev-statistics-010814-020148.
Straus, S.G., Lewis, M.W., Connor, K., Eden, R., Boyer, M.E., Marler, T., Carson, C.M., Grimm, G.E., Smigowski, H., 2019. Collective simulation-based training in the US Army: User interface fidelity, costs, and training effectiveness. RAND Corporation, Santa Monica, CA. https://doi.org/10.7249/RR2250.
Sutherland, W.J., Woodroof, H.J., 2009. The need for environmental horizon scanning. Trends Ecol. Evol. 24, 523–527. https://doi.org/10.1016/j.tree.2009.04.008.
Sutton, R.S., Barto, A.G., 2018. Reinforcement learning: An introduction, 2nd ed. MIT press, Cambridge, Massachusetts. ISBN: 978-026203924.
Taylor, N., 2003. Review of the use of models in informing disease control policy development and adjustment, Report to DEFRA. University of Reading, Reading, United Kingdom, 98 p. ⟨http://www.veeru.rdg.ac.uk/documents/UseofModelsinDiseaseControlPolicy.pdf⟩.
Tildesley, M.J., Savill, N.J., Shaw, D.J., Deardon, R., Brooks, S.P., Woolhouse, M.E., Grenfell, B.T., Keeling, M.J., 2006. Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. Nature 440, 83–86. https://doi.org/10.1038/nature04324.
Truelove, S., Smith, C.P., Qin, M., Mullany, L.C., Borchering, R.K., Lessler, J., Shea, K., Howerton, E., Contamin, L., Levander, J., Kerr, J., Hochheiser, H., Kinsey, M., Tallaksen, K., Wilson, S., Shin, L., Rainwater-Lovett, K., Lemairtre, J.C., Dent, J., Kaminsky, J., Lee, E.C., Perez-Saez, J., Hill, A., Karlen, D., Chinazzi, M., Davis, J.T., Mu, K., Xiong, X., Pastore y Piontti, A., Vespignani, A., Srivastava, A., Porebski, P., Venkatramanan, S., Adiga, A., Lewis, B., Klahn, B., Outten, J., Orr, M., Harrison, G., Hurt, B., Chen, J., Vullikanti, A., Marathe, M., Hoops, S., Bhattacharya, P., Machi, D., Chen, S., Paul, R., Janies, D., Thill, J.-C., Galanti, M., Yamana, T.K., Pei, S., Shaman, J.L., Healy, J.M., Slayton, R.B., Biggerstaff, M., Johansson, M.A., Runge, M.C., Viboud, C., 2022. Projected resurgence of COVID-19 in the United States in July—December 2021 resulting from the increased transmissibility of the Delta variant and faltering vaccination. eLife 11, e73584. https://doi.org/10.7554/eLife.73584.
Tyszkiewicz, M.T., McCleskey, E.R., Miller, R., 2012. Updating the national planning scenarios: using wicked problems and capability-based planning concepts for homeland security. J. Homel. Secur. Emerg. Manag. 9 https://doi.org/10.1515/1547-7355.1903.
Wald, A., 1947. Sequential analysis. John Wiley & Sons, New York, NY.
Wade-Malone, L.K., Howerton, E., Probert, W.J.M., Runge, M.C., Viboud, C., Shea, K., 2024. When do we need multiple infectious disease models? Agreement between projection rank and magnitude in a multi-model setting. Epidemics 47, 100767. https://doi.org/10.1016/j.epidem.2024.100767.
Walker, P., Whittaker, C., Watson, O., Baguelin, M., Ainslie, K., Bhatia, S., Bhatt, S., Boonyasiri, A., Boyd, O., Cattarino, L., Cucunuba, Z., Cuomo-Dannenburg, G., Dighe, A., Donnelly, C.A., Dorigatti, I., van Elsland, S., others, 2020. Report 12: The global impact of COVID-19 and strategies for mitigation and suppression. Imperial College COVID-19 Response Team, London, United Kingdom, p. 18. https://doi.org/10.25561/77735.
Walters, C.J., 1986. Adaptive management of renewable resources. Macmillan, New York, New York, USA.
Whitehead, M., 1992. The concepts and principles of equity and health. Int. J. Health Serv. 22, 429–445. https://doi.org/10.2190/986L-LHQ6-2VTE-YRRN.