Choosing the Appropriate Modeling Method for a Given Problem: Health Economic Modeling, Causal Modeling, Simulation, or Constrained Optimization?

William Crown, MA, PhD, OptumLabs, Cambridge, MA, USA; **Kalyan S. Pasupathy, PhD**, Mayo Clinic, Rochester, MN, USA; **Deborah A. Marshall, PhD**, University of Calgary, Calgary, AB, Canada; **Praveen Thokala, PhD**, University of Sheffield, Sheffield, England, UK

Massive expansion in the availability of data combined with advances in analytic methods create tremendous opportunities for HEOR analyses. But multidisciplinary teams will be necessary to realize these opportunities.

There are a variety of analytic methods available to researchers for approaching different types of health economic evaluation problems. Most researchers have expertise in a specific analytic method such as health economic modeling or causal inference from health econometrics/epidemiology. More recently, we are seeing an increased use of constrained optimization and simulation methods. These methods are often highly complementary, but analytic opportunities are lost because deep methodological domain knowledge keeps researchers locked within their own methodological silos. For example, discrete event simulation methods are widely used in health economic modeling, and causal modeling methods are often a precursor to estimate the parameters in health economic models or building the equations in a simulation model.

In this article, we consider 4 major analytic methods: (1) health economic modeling, (2) causal modeling, (3) simulation modeling, and (4) constrained optimization modeling. We propose that the complementarity of the insights produced by the different methods argues for the benefits of building interdisciplinary teams of researchers with different methodological skillsets.

Health Economic Modeling: Building the Patient Footprint

Health economic modeling is widely applied in cost-effectiveness evaluations of pharmaceutical products, devices, and other interventions by health technology assessment organizations and payers to assess the value of new treatments.[1] Why do we need modeling? One important reason is that the data necessary to conduct cost-effectiveness analyses typically reside in different places and must be combined using a modeling framework. As indicated in Figure 1, many different inputs are needed for health economic models. These include treatment effectiveness, cost and resource use, quality of life, and adverse events. For example, health technology assessment organizations typically evaluate new technologies following marketing approval by regulatory authorities. The primary information available at the time of approval is the efficacy and safety evidence from the randomized controlled trials used for the regulatory submission. Since there is no market evidence based on experience with the product yet, the cost and patient utility data must be gathered from other sources for similar patient populations. It is also important to understand the natural history of disease for the condition being evaluated, and it is necessary to understand the quality of the data sources for each of these inputs. Due to the maturity of the health economics modeling field, there are many guidelines for building health economic models.

Figure 1. Inputs for health economic models.

TYPES:	SOURCES:	USES:
Effectiveness	“Published papers”	Parameter values
Costs	Routine data	Model structure
Resource use/activity	Reference sources	Sensitivity analysis
Health states	Local/clinical/expert opinion	Validation/consistency/calibration
Utility values	Sponsor submissions	
Indirect comparators		
Longer-term outcomes		
“Other” interventions		
Natural history		
Epidemiology		

Causal Modeling: Estimating the Impact of an Intervention

The strongest causal inferences come from randomized designs that balance interventions on both observable and unobservable confounders. Randomized designs also greatly simplify the statistical analysis of treatment effects. However, for many reasons, evidence from randomized trials often is not available. As a result, researchers attempt to draw causal inferences from secondary data sources not originally intended to support research. After a product has been on the market long enough, evidence on a product begins to accumulate in medical claims and electronic health records. We have good statistical methods for addressing many of the issues that arise in the analysis of observational data. However, in observational analyses, we need to be careful about design and statistical methods in order to arrive at reliable inferences.[2] The methods from epidemiology—propensity score, inverse probability weights, G-estimation, and so forth—are extremely important, but the most important contribution from epidemiologists is what they've taught us about research design. Economists have developed a complementary set of methods that use empirical correlations in the error structures of models to correct for a wide variety of measurement and specification challenges common in real-world data analysis. These include parametric and nonparametric sample selection bias models, as well as a broad range of simultaneous equations methods.

Simulation Models: Analyzing Complex Systems

Simulation models use the results from causal models and health economic models to evaluate problems from a systems perspective.[3] This requires thinking about the context (including the people, technology, and healthcare settings) in which these services and technologies are delivered. Healthcare delivery processes include feedback loops, as well as nonlinear and spatial relationships among entities, multiple agents or stakeholders, time dependency and dynamic transitions within the system, and the idea of emergency. "Emergency" is not used in the context of being urgent, but rather how things emerge downstream, resulting in intended and unintended consequences in the system. For example, it is very difficult to anticipate how patients will interact with the healthcare system, and how this will affect individual patient outcomes and health system performance outcomes (eg, wait times). The key idea around simulation modeling is to model the complexity of the system, and then evaluate results for various "what if" scenarios to inform planning for healthcare services delivery. Importantly, simulation enables assessment not only of intended effects but also unintended effects that may not be anticipated due to system complexity. Using simulation modeling makes it possible to explore and anticipate the impact of potential changes without actually altering the system until a strategy or policy has been identified that improves overall system performance.

> The COVID-19 pandemic sweeping the globe provides a poignant example of how the 4 methods can be applied to address different components of a critical problem.

Constrained Optimization: Using Math to Set Policy

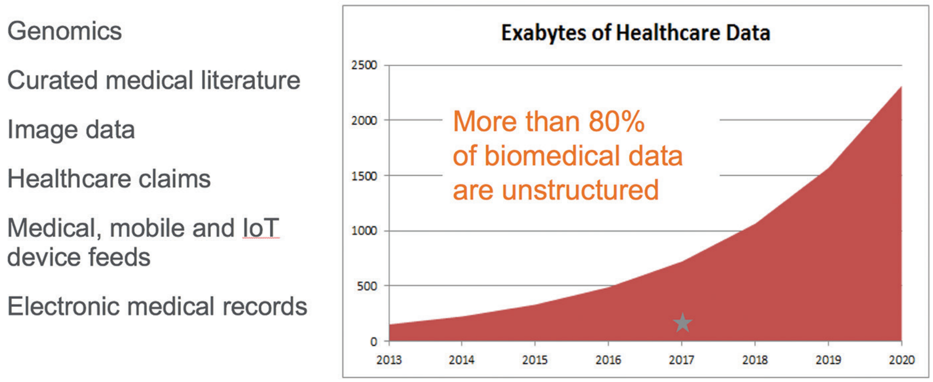
A fourth methodological approach is constrained optimization. The term "optimal" is widely and loosely used in healthcare. Constrained optimization is a mathematical approach to finding the truly best solution to a problem, subject to real-world constraints.[4] In health technology assessment analyses, for example, we can use constrained optimization to identify the most cost-effective policy decision subject to real-world constraints such as the health system budget. Constrained optimization methods are a tool for dealing with the combinatorial complexity of healthcare problems that overwhelm decision makers leading them to make suboptimal decisions. They consist of an objective function that we are trying to optimize (eg, minimize the number of cervical cancer cases), a set of decision or policy variables (eg, cervical cancer screening or vaccination for human papilloma virus), a set of parameters for each of the decision variables (these are externally determined prior to the optimization modeling), and a set of constraints (eg, budget constraint). As with each of the other methods, there are many different types of constrained optimization modeling approaches, depending upon the problem.

Matching Methods to Problems: the COVID-19 Pandemic

In this overview, we have briefly summarized 4 major types of methods: health economic modeling, causal modeling, simulation modeling, and optimization modeling. And although there are many different methods that are used in health economics and outcomes research, it's probably fair to say that most fall within these 4 major types of methods.

The COVID-19 pandemic sweeping the globe provides a poignant example of how the 4 methods can be applied to address different components of a critical problem. The nonlinearity of disease transmission, the differential mortality among alternative population subgroups, and the differential supply of medical services across geographies all render the traditional methods used by health systems inadequate to anticipate where critical shortfalls in needed care may occur. This is a problem that is tailor-made for simulation models. SIR models from epidemiology are systems of differential equations that model the population susceptible, infected, or recovered (or, alternatively, removed).[5] The parameters in the model are calibrated for local characteristics and enable "what if" simulations in response to changes in assumptions. Agent-based simulation models can extend SIR models to include agents interacting with different groups in the community such as schools, places of employment, grocery stores, or the healthcare system. Similarly, one could use discrete event simulation to estimate the demand for specific types of healthcare services that could then be evaluated given the level of local supply (eg, number of hospital beds, ventilators, nurses, and physicians) available through real-world data analyses. After the first

Figure 2. Healthcare Big Data.



https://www.cio.com/article/2860072/healthcare/how-cios-can-prepare-for-healthcare-data-tsunami.html

wave of the pandemic has passed, a tremendous amount of data will have been generated on how patients were treated. These data reflect a series of natural experiments that enable the performance of alternative treatment approaches to be assessed using causal inference methods. Similarly, the cost-effectiveness of these alternative treatment approaches can be assessed using health economic modeling. Finally, assuming that some of the existing therapies used to treat COVID-19 patients were shown to be effective, or newly developed therapies have become available, constrained optimization methods could be used to design optimal screening and treatment protocols. This has already been done successfully for the treatment of influenza.[6] In short, it is likely that all 4 categories of models will be highly relevant for dealing with the COVID-19 pandemic and preparing us for subsequent waves of the virus.

The COVID-19 example illustrates that there are multiple factors that play into selecting an analytic approach to a problem. Rarely are the methods mutually exclusive, and they are often highly complementary. The example clearly illustrates the value of considering an expanded selection of methods that may help frame a more complete solution than might be possible by staying within a particular methodological silo. To do so, however, requires an expanded skill set. ISPOR members are generally familiar with health economic modeling and the causal modeling methods from epidemiology, econometrics, and health services research. However, the skill sets needed for simulation and optimization relate to the field of operations research that has traditionally been the bastion of engineering. (Although it is clear from their use of SIR models that mathematical epidemiologists have been working with simulation methods for many years!)

What's Next for HEOR Models?

Looking ahead, machine learning is yet another method that is coming to us from engineering and computer science.[7,8] We are starting to see a need for teams with training in economics, epidemiology, engineering, and computer science as we move into this new environment where we have access to much more data—much of which are unstructured (Figure 2). In addition, healthcare domain knowledge is very important to augment the technical skills of the various types of modelers. Those trained solely in machine learning methods often lack experience with observational data and knowledge of the healthcare sector. Conversely, those trained in epidemiology, health economics, and health services research generally lack skills in natural language processing and machine learning techniques that will be needed to deal with unstructured data, complex data structures, and data volume that are already with us today. The health economics and outcomes research challenges of the future will require us to move beyond our methodological silos and build multidisciplinary teams. •

References

1. Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS)—Explanation and elaboration: a report of the ISPOR health economic evaluations publication guidelines good reporting practices task force. *Value Health*. 2013;16(2):231-250.

2. Johnson M, Crown W, Martin B, Dormuth C, Siebert U. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from non-randomized studies of treatment effects using secondary data sources. Report of the ISPOR retrospective database analysis task force—Part III. *Value Health*. 2009;12(8):1062-1073.

3. Marshall D, Ijzerman M, Crown W, et al. Applying dynamic simulation modeling methods in health care delivery research: the SIMULATE checklist. *Value Health*. 2015;18(1):5-16.

4. Crown W, Buyukkaramikli N, Thokala P, et al. Constrained optimization methods in health services research—an introduction: report 1 of the ISPOR optimization methods emerging good practices task force. *Value Health*. 2017; 20(3):310-319.

5. Brauer F, van den Driessche P, Wu J. *Mathematical Epidemiology*. Berlin, Germany: Springer;2008

6. Standaert B, Van Vlaenderen I, Van Bellinghen LA, et al. Constrained optimization for the selection of influenza vaccines to maximize the population benefit: a demonstration project. *Appl Health Econ Health Policy*. 2019;(Epub 22 November 2019).

7. Crown W. Real-world evidence, causal inference, and machine learning. *Value Health*. 2019;22(5):587-592.

8. Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. *Value Health*. 2019;(22)7:808-815.