

Book website (with statistical code for each chapter):

<https://github.com/sanjaybasu/modelinghealthsystems>

Chapter 1:

In this Chapter, we define and provide examples of several key terms used in public health and healthcare modeling research. We begin by clarifying the differences between key terms used to describe rates of disease (incidence, prevalence, and mortality) as well as the performance characteristics of tests used to detect disease (sensitivity, specificity, positive predictive value, and negative predictive value), prevent or treat disease (odds ratios, relative risks), understand studies (case-control, cohort, and randomized controlled trials), and avoid common study problems (bias, confounding).

Chapter 2:

In this Chapter, we seek to answer the question: how much should we pay for a public health program? We often have to decide how to allocate funds to different public health programs, or decide whether a new medical test or treatment is worth the cost. How can we make such decisions fairly? We'll first work through some examples of decision trees that are commonly used to make these judgments in a rigorous and fair way. We'll create some decision trees to perform cost-benefit analysis to determine whether we want to pay for a new service, test, or treatment if we are focused on lowering the costs of our operations. We'll then understand how to perform cost-effectiveness analysis to identify under what circumstances a more expensive new service, test, or treatment might be worth the cost because it meaningfully improves health outcomes.

Chapter 3:

In this Chapter, we will examine one of the most common and useful ways to understand complex public health and healthcare interventions: the Markov model. A Markov model is a representation of health or disease in multiple stages. In this Chapter, we will use Markov models to identify how effective or cost-effective our public health and healthcare programs might be. Markov models are highly flexible and allow for an infinite variety of diseases or interventions to be simulated and understood, which is why they are among the most popular tools for public health and healthcare research.

Chapter 4:

In this Chapter, we introduce the statistical program R, which provides a straightforward approach to modeling, and has the advantages of being free, fast, available on any operating system, commonly-used, and widely-supported by an online community that helps users who get stuck. We provide a detailed introduction to using R that will be useful for the more advanced modeling methods we introduce and practice in the next several Chapters.

Chapter 5:

In previous Chapters, we ignored a critical aspect of modeling some major diseases: the infectious nature of many diseases. For infectious diseases, the risk of getting the disease is related to how many people are infectious at a given time; the more infectious people in the area, the higher the risk of infection among susceptible people. In a typical Markov model, we can't account for this basic feature of infectious diseases, because the risk of moving from one state (healthy) to another state (diseased) is assumed to be constant. In Chapter 9, we introduce a simulation modeling framework that has been used for decades to simulate infectious disease epidemics.

Chapter 6:

In the next Chapter, we defined and simulated the most basic epidemic model. We assumed that

in the prior Chapter, we derived and simulated the most basic epidemic model. We assumed that people can be in only one of three states (susceptible, infected, or recovered), and that people mix homogeneously throughout the population. In Chapter 10, we examine how the Kermack-McKendrick model can be extended to simulate a wide variety of complex diseases and circumstances, and be adapted to incorporate the complex ways that people contact each other. We additionally describe methods for simulating individual behavior in response to an epidemic.

Chapter 7:

In previous Chapters, we used Markov models to estimate the burden of disease, and the potential impact of our interventions. One of the key limitations to Markov models is that they don't take into account a person's unique individual characteristics; Markov models are designed to efficiently simulate the average outcome for an entire population. For many public health and healthcare system problems, however, we need to consider heterogeneity within a population, or differences in risk and differences in benefit from our programs. For that purpose, microsimulation models, which take into account unique characteristics of individuals and the correlations between these characteristics, can be more useful. In Chapter 8, we detail the constructing and use of microsimulation models, using examples related to diabetes prevention and treatment.

Chapter 8:

In this Chapter, we use the power of microsimulations to identify how to better personalize medical treatment. We describe the concept of heterogeneous treatment effects, which describe how one group of people might have a systematically different outcome (benefit or risk) from a treatment than another group. We describe how to compare clinical guidelines in this setting, and identify the best treatment for an individual patient by combining models with randomized trial data.

Chapter 9:

In this Chapter, we focus on big data analyses and the implementation of machine learning methods to better identify how much one intervention might affect different people in the population differently. Extending the concept of heterogeneous treatment effects from the prior chapter, we examine several strategies for implementing common machine learning methods to better classify and predict disease: elastic net regularization, random forest methods, gradient boosting machines, and deep learning neural networks.

Chapter 10:

In this Chapter, we extend our work to understand the concept of validating models against real-world data. We specifically discuss the study of 'natural experiments', or rare opportunities where real-world data can help us identify whether a models' predictions are accurate. We focus on two common methods of natural experiments research—regression discontinuity and difference-in-differences analysis—to evaluate how much impact an intervention can have on different groups of people.

Chapter 11:

Throughout this book, we have focused on the practices of constructing our own models or using standard modeling templates and strategies to solve common public health and healthcare system problems. But inherent to the task of using models is the challenge of being a good consumer of models. Often, we are faced with the task of reading and interpreting models produced by others, and determining whether we "believe" the model results and can make use of the model implementation to help us make decisions. In this Chapter, we address the issue of how we might become better consumers of modeling studies.

