Preface

Why This Book?

This book is motivated by the following convictions:

1) Quantitative risk assessment (QRA) can be a powerful discipline for improving risk management decisions and policies.
2) Poorly conducted QRAs can produce results and recommendations that are worse than useless.
3) Sound risk assessment methods provide the benefits of QRA modeling – being able to predict and compare the probable consequences of alternative actions, interventions, or policies and being able to identify those that make preferred consequences more probable – while avoiding the pitfalls.

This book develops and illustrates QRA methods for complex and uncertain biological, engineering, and social systems. These systems have behaviors that are too complex or uncertain to be modeled accurately in detail with high confidence. Practical applications include assessing and managing risks from chemical carcinogens, antibiotic resistance, mad cow disease, terrorist attacks, and accidental or deliberate failures in telecommunications network infrastructure.

For Whom Is It Meant?

This book is intended primarily for practitioners who want to use rational quantitative risk analysis to support and improve risk management decisions in important health, safety, environmental, reliability, and security applications, but who have been frustrated in trying to apply traditional quantitative modeling methods by the high uncertainty and/or complexity of the systems involved. We emphasize methods and strategies for modeling causal relations in complex and uncertain systems well enough to make effective risk management decisions. The book is written for practitioners from multiple disciplines – decision and risk analysts, operations researchers and management scientists, quantitative policy analysts, economists, health and

safety risk assessors, engineers, and modelers – who need practical ways to predict and manage risks in complex and uncertain systems.

What's in It?

Three introductory chapters describe QRA and compare it to less formal alternatives, such as taking prompt action to address current concerns, even if the consequences caused by the recommended action are unknown (Chapter 1). These chapters survey QRA methods for engineering risks (Chapter 2) and health risks (Chapter 3). Brief examples of applications such as flood control, software failures, chemical releases, and food safety illustrate the scope and capabilities of QRA for complex and uncertain systems.

Chapter 1 discusses a concept of *concern-driven risk management*, in which qualitative expert judgments about whether concerns warrant specified risk management interventions are used in preference to QRA to guide risk management decisions. Where QRA emphasizes the formal quantitative assessment and comparison of the probable consequences caused by recommended actions to the probable consequences of alternatives, including the status quo, concern-driven risk management instead emphasizes the perceived urgency or severity of the situation motivating recommended interventions. In many instances, especially those involving applications of a "Precautionary Principle" (popular in much European legislation), no formal quantification or comparison of probable consequences for alternative decisions is seen as being necessary (or, perhaps, possible or desirable) before implementing risk management measures that are intended to prevent serious or irreversible harm, even if the causal relations between the recommended measures and their probable consequences are unclear. Such concern-driven risk management has been recommended by critics of QRA in several areas of applied risk management.

Based on case studies and psychological literature on the empirical performance of judgment-based decision making under risk and uncertainty, we conclude that, although concern-driven risk management has several important potential political and psychological advantages over QRA, it often performs less well than QRA in identifying risk management interventions that successfully protect human health or achieve other desired consequences. Therefore, those who advocate replacing QRA with concern-driven alternatives, such as expert judgment and consensus decision processes, should assess whether their recommended alternatives truly outperform QRA, by the criterion of producing preferred consequences, before rejecting the QRA paradigm for practical applications.

Chapter 2 introduces methods of probabilistic risk assessment (PRA) for predicting and managing risks in complex engineered systems. It surveys methods for PRA and decision making in engineered systems, emphasizing progress in methods for dealing with uncertainties, communicating results effectively, and using the results to guide improved decision making by multiple parties. For systems operating under threats from intelligent adversaries, novel methods and game-theoretic ideas can

help to identify effective risk reduction strategies and resource allocations. In hard decision problems, where the best course of action is unclear and data are sparse, ambiguous, or conflicting, state-of-the-art methodology can be critical for good risk management. This chapter discusses some of the most useful PRA methods and possible extensions and improvements.

Chapter 3 introduces methods of quantitative risk assessment (QRA) for public health risks. These arise from the operation of complex engineering, economic, medical, and social systems, ranging from food supply networks to industrial plants to administration of school vaccination programs and hospital infection control programs. The decisions and behaviors of multiple economic agents (e.g., the producers, distributors, retailers, and consumers of a product) or other decision makers (e.g., parents, physicians, and schools involved in vaccination programs) affect risks that, in turn, typically affect many other people. Health risks are commonly different for different subpopulations (e.g., infants, the elderly, and the immunocompromised, for a microbial hazard; or customers, employees, and neighbors of a production process). Thus, public health risk analysis often falls in the intersection of politics, business, law, economics, ethics, science, and technology, with different participants and stakeholders favoring different risk management alternatives. In this politicized context, QRA seeks to clarify the probable consequences of different risk management decisions.

Chapters 4 and 5 (as well as Chapter 15, which deals specifically with terrorism risk assessment) emphasize that *sound risk assessment requires developing sound risk models* in enough detail to represent correctly the (often probabilistic) *causal relations* between a system's controllable inputs and the outputs or consequences that decision makers care about. "Sound" does not imply completely accurate, certain, or detailed. Imperfect and high-level risk models, or sets of alternative risk models that are contingent on explicitly stated assumptions, can still be sound and useful for improving decision making. But a sound model must describe causal relations correctly, even if not in great detail, and even if contingent on stated assumptions. Incorrect causal models, or models with hidden false assumptions about cause and effect, can lead to poor risk management recommendations and decisions.

Chapters 4 and 5 warn against popular shortcut methods of risk analysis that try to avoid the work required to develop and validate sound risk models. These include replacing empirically estimated and validated causal risk models (e.g., simulation models) with much simpler ratings of risky prospects using terms such as *high*, *medium*, and *low* for attributes such as the frequency and severity of adverse consequences. Other shortcut methods use highly aggregate risk models or scoring formulas (such as "*risk = potency × exposure,*" or "*risk = threat × vulnerability × consequence*") in place of more detailed causal models. Many professional consultants, risk assessors, and regulatory agencies use such methods today. However, *these attempted shortcuts do not work well in general.* As discussed in Chapters 4 and 5, they can produce results, recommendations, and priorities that are worse than useless: they are even less effective, on average, than making decisions randomly! Poor risk management decisions, based on false predictions and assumptions, result from these shortcut methods.

Fortunately, it is possible to do much better. Building and validating sound causal risk models leads to QRA models and analyses that can greatly improve risk management decisions. Chapters 6 through 16 explain how. They introduce and illustrate techniques for testing causal hypotheses and for identifying potential causal relations from data (Chapters 6 and 7), for developing (and empirically testing and validating) risk models to predict the responses of complex, uncertain, and nonlinear systems to changes in controllable inputs (Chapters 8-13), and for making more effective risk management decisions, despite uncertainties and complexities (Chapters 14-16). These chapters pose a variety of important risk analysis challenges for complex and uncertain systems, and propose and illustrate methods for solving them in important real-world applications.

Key challenges, methods and applications in Chapters 6 through 16 include the following:

- *Information-theory and data-mining algorithms.* Chapter 6 shows how to detect initially unknown, possibly nonlinear (including *u*-shaped) causal relations in epidemiological data sets, using food poisoning data as an example. A combination of information theory and nonparametric modeling methods (especially, classification tree algorithms) provide constructive ways to identify potential causal relations (including nonlinear and multivariate ones with high-order interactions) in multivariate epidemiological data sets.
- *Testing causal hypotheses and discovering causal relations.* Chapter 7, building on the methods in Chapter 6, discusses how to test causal hypotheses using data, how to discover new causal relations directly from data without any a priori hypotheses, and how to use data mining and other statistical methods to avoid imposing one's own prior beliefs on the interpretation of data – a perennial challenge in risk assessment and other quantitative modeling disciplines. An application to antibiotic-resistant bacterial infections illustrates these techniques.
- *Use of new molecular-biological and "-omics" information in risk assessment.* Chapter 8 shows how to use detailed biological data (arising from advances in genomics, proteomics, metabolomics, and other low-level biological data) to predict the fraction of illnesses, diseases, or other unwanted effects in a population that could be prevented by removing specific hazards or sources of exposure. This challenge is addressed by using conditional probability formulas and conservative upper bounds on the observed occurrence and co-occurrence rates of events in a causal network to obtain useful upper bounds on unknown causal fractions. Bounding calculations are illustrated by quantifying the preventable fraction of smoking-associated lung cancers in smokers caused by – and preventable by blocking – a particular causal pathway (involving polycyclic aromatic hydrocarbons forming adducts with DNA in a critical tumor suppressor gene) that has attracted great recent interest.
- *Upper-bounding methods.* Chapters 8 through 12 consider how to use available knowledge and information about causal pathways in complex systems, even if very imperfect and incomplete (e.g., biomarker data for complex diseases), to estimate upper bounds on the *preventable fractions* of disease that could be

eliminated by removing specific hazardous exposures. (Analogous strategies for using partial information to bound the preventable risks of adverse outcomes can be used for other complex systems.) The applications in these chapters focus on antibiotic-resistant bacterial infections and on smoking-related lung cancers as examples of partly understood complex systems with large and important knowledge and data gaps, but with enough available knowledge about causal pathways to be useful.

- *Identification of a discrete set of possible risks.* Using dose-response relations for lung cancer risk as an extended example, Chapters 10 and 11 show how to quantify several different input-output relations for a complex system that are consistent with available knowledge and data about uncertain causal mechanisms. Chapter 10 addresses how to identify promising leads for R&D on designing a less hazardous cigarette. It uses a portfolio of causal mechanisms to identify removing cadmium as a promising (but uncertain) way to reduce total risk, despite the complexity of the mixture of chemicals to which smokers are exposed, the complex and uncertain biological pathways by which these chemicals affect lung cancer risk, and the many scientific uncertainties that remain. Chapter 11 shows that sometimes the response of a complex system to a change in inputs can be identified as one of a small number of equally probable alternatives, all of which are consistent with past data.
- *Systems dynamics analysis and simulation.* Chapters 10 through 13 illustrate how to predict input-output relations of dynamic systems using simulation modeling and mathematical analysis (solution of systems of ordinary differential equations and algebraic equations), derived from empirical data and knowledge of the causal processes being simulated. Systems dynamics models can benefit from other techniques demonstrated in these chapters, including modeling only the steady-state levels of subprocesses that adjust relatively quickly and that affect slower processes primarily through time-averaged values (so that hard-to-model but brief, bounded transients can safely be ignored) and using Markov's inequality to relate deterministic simulations of mean values to bounds on probable values of underlying stochastic processes.
- *Comparative statics analysis and reduction of complex models*. Chapter 13 discusses how to reduce large dynamic models, represented by networks of interacting dynamic processes, to much smaller ones that predict the same equilibrium behaviors in response to changes in inputs.
- *Decision tree, sequential decision optimization, and value of information (VOI) analysis*. Chapter 14 estimates the economic value of information from tracking country-of-origin information for cattle imported into the United States from Canada (or other countries with "mad cow" disease). Deliberately using worse-than-realistic probability distributions for scenarios yields a lower bound on the economic value of information (VOI) from tracking. [The author has long believed that the USDA's policy of allowing Canadian cattle – especially, older cattle – into the United States is inconsistent with the policy goal of keeping mad cow disease (bovine spongiform encephalitis, BSE) out of the United States; he has served as an expert in litigation intended to force the USDA to reconsider and

revise this policy.] Assuming that the USDA continues to allow these imports, Chapter 14 considers how to manage the resulting economic risks to the United States created by the increased probability that another case of BSE in an animal imported from Canada will be discovered. The analytic methods demonstrated in Chapter 14 are also useful for many other public risk management and policy optimization applications in which future events and decisions affect the eventual outcomes of present decisions.

- *Game-theory and hierarchical optimization models*. Modeling the behaviors of intelligent attackers and intelligent defenders of a facility (or other target) and optimizing the allocation of defensive resources, taking into account how attackers may respond, are crucial topics in terrorism risk analysis. Methods currently in widespread use for these challenges have serious limitations, and improved methods are urgently needed. Chapter 15 considers both the limitations and ways to improve upon current methods of terrorism risk analysis.
- *Mathematical optimization and phase-transition modeling*. Chapter 16 surveys methods for predicting the resilience of complex systems (e.g., telecommunications networks) to deliberate attacks, and for designing systems to make them resilient to attack. One of the key ideas in this chapter is that the dynamic behaviors of large networks can be extremely simple. For example, simple statistical ("scale-free") models of telecommunications networks predict almost complete resilience to attacks that are limited to knocking out at most a small number of nodes (or links) simultaneously, provided that each node has "enough" (at least a certain critical percentage) of surplus routing capacity to handle the displaced traffic. (Here, "resilient" means that at most only a small fraction of traffic between other nodes, approaching zero percent in large networks, will be made unroutable by such an attack.) At the same time, these simple models predict that networks may be highly vulnerable to such attacks (meaning that most of the traffic in the network will become unroutable after the initial attacks cause node overloads and failures to cascade through the network) if each node has less than the critical amount of surplus capacity. Such a "phase transition" (with a transition threshold determined by the critical amount of surplus capacity) from resilient to vulnerable is characteristic of many highly idealized models of scale-free networks. Assuming that real networks have similar phase-transition behavior – which is currently an important unknown – individual network owners and operators may still lack incentives to invest in increasing resilience, even if doing so would benefit them collectively.

Some Specific Risk Models and Applications for Interested Specialists

In addition to general risk modeling methods, several chapters present specific risk models and results that may be of independent interest to scientists and researchers in cancer risk analysis, bioinformatics and toxicology, microbial and antimicrobial

risk assessment, food safety, and terrorism risk analysis. For example, Chapters 11 and 12 develop and apply a new model of lung carcinogenesis. Exposure-related carcinogenesis is often modeled by assuming that cells progress between successive stages – possibly undergoing proliferation at some of them – at rates that depend (usually linearly) on biologically effective doses. Biologically effective doses, in turn, may depend nonlinearly on administered doses, due to pharmacokinetic nonlinearities. Chapter 11 provides a mathematical analysis of the expected number of cells in the last ("malignant") stage of a "multistage clonal expansion" (MSCE) model as a function of dose rate and age. The solution displays symmetries such that several distinct sets of parameter values fit past epidemiological data equally well. These different possible sets of parameter values make identical predictions about how changing exposure levels or timing would affect risk. Yet they make significantly different predictions about how changing the composition of exposure would affect risk. Biological data, revealing which rate parameters describe which specific stages, are required to yield unambiguous predictions. From epidemiological data alone, only a set of equally likely alternative predictions can be made for the effects on risk of such interventions.

Chapter 12 asks the following question: If a specific biological mechanism could be discovered by which a carcinogen increases lung cancer risk, how might this knowledge be used to improve risk assessment? For example, suppose that arsenic in cigarette smoke increases lung cancer risk by hypermethylating the promoter region of a specific gene (p16INK4a), leading to more rapid entry of altered (initiated) cells into a clonal expansion phase. How could the potential impact on lung cancer of removing arsenic be quantified in light of such knowledge (assuming, for purposes of illustration, that this proposed mechanism is correct)? Chapter 12 provides an answer, using a three-stage version of the MSCE model from Chapter 11. [This refines a more usual two-stage clonal expansion (TSCE) model of carcinogenesis by resolving its intermediate or "initiated" cell compartment into two subcompartments, representing experimentally observed "patch" and "field" cells. This refinement allows p16 methylation effects to be represented as speeding transitions of cells from the patch state to the clonally expanding field state.] Given these assumptions, removing arsenic might greatly reduce the number of non-small cell lung cancer cells produced in smokers, by up to two thirds, depending on the fraction (between 0 and 1) of the smoking-induced increase in the patch-to-field transition rate prevented if arsenic were removed. At present, this fraction is unknown (and could be as low as zero), but the possibility that it could be high (close to 1) cannot be ruled out without further data.

Chapter 13 presents a dynamic disease model for chronic obstructive pulmonary disease (COPD), a family of smoking-associated diseases having complex causes and consequences. It shows how improved understanding of interactions among biological processes, and of how exposures (in this case, to cigarette smoke) affect these processes and their interactions, can be used to better predict health risks caused by exposures. COPD, although the fourth-leading cause of death worldwide, has a puzzling etiology. It is a smoking-associated disease, but only a minority of smokers develop it. Moreover, in people (but not in animals), unresolved inflammation of the

lung and destruction of lung tissue, once started, continue even after smoking ceases. Chapter 13 proposes a biologically based risk assessment model of COPD that offers a possible explanation of these and other features of the disease. COPD causation is modeled as resulting from a dynamic imbalance between protein-digesting enzymes (proteases) and the antiproteases that inhibit them in the lung. This leads to ongoing proteolysis (digestion) of lung tissue by excess proteases. The model is formulated as a system of seven ordinary differential equations (ODEs) with 18 parameters to describe the network of interacting homeostatic processes regulating the levels of key proteases and antiproteases. Mathematical analysis shows that this system can be simplified to a single quadratic equation to predict the equilibrium behavior of the entire network. There are two possible equilibrium behaviors: a unique stable "normal" (healthy) equilibrium, or a "COPD" equilibrium with elevated levels of lung macrophages and neutrophils (and their elastases) and reduced levels of antiproteases. The COPD equilibrium is induced only if cigarette smoking increases the average production of macrophage elastase (MMP-12) per alveolar macrophage above a certain threshold. Following smoking cessation, the COPD equilibrium levels of MMP-12 and other disease markers decline but do not return to their original (presmoking) levels. These and other predictions of the model are consistent with limited available human data.

Chapters 14, 15, and 16 present risk models for systems in which the future decisions of multiple participants affect the final consequences of current decisions. These chapters present several example models and results for "mad cow" disease (BSE) risk management, terrorist risk analysis, and risk analysis of telecommunications network infrastructure.

Why Do These Models and Methods Matter?

The main purpose of the specific models and applications in the later chapters, as well as of the general QRA methods in earlier chapters, is to show how *QRA can be carried out successfully for uncertain, complex, and nonlinear systems of great practical importance.* Some skeptics have argued that QRA modeling is impractical and/or too laden with uncertain assumptions to give useful and trustworthy results in practice (see Chapter 1). This book seeks to show, both through general modeling principles and by means of constructive examples, how QRA can successfully be carried out and used today to improve risk management in a variety of important real-world applications.

http://www.springer.com/978-0-387-89013-5

Risk Analysis of Complex and Uncertain Systems

Cox Jr., L.A.

2009, XXVIII, 436 p., Hardcover
ISBN: 978-0-387-89013-5